

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representation of
The original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 ; Search time 9.81254 Seconds

(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-19

Perfect score: 218

Sequence: 1 cgcacagaaagttatttca.....caccagtgttgcacattca 714

Scoring table:

OLIGO	
Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 262574 seqs, 29422322 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09835992/Runat14012003.161527.4050/app_query.fasta.1.3932
-DB=Issued_Patents_AA -OPMT=fastan -SUFFIX=olig.ra1 -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFM=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09835992.cgn2_1.1_40_@runat14012003.161527.4050 -NCPU=6 -ICPU=3
-NO_XLPLY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Issued_Patents_AA:
1: /cgn2_6/plodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/plodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/plodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/plodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/plodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/plodata/1/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3.2	18	6	5182195-2	Patent No. 5182195
2	3.2	19	4	US-09-149-476-534	Sequence 534, App
3	3.2	21	6	5182195-7	Patent No. 5182195
4	3.2	44	4	US-09-288-143-88	Sequence 88, Appl
5	3.2	62	1	US-08-488-961-6	Sequence 6, Appl1
6	3.2	62	4	US-08-973-297-6	Sequence 6, Appl1
7	3.2	62	5	PCT-US96-06511-6	Sequence 6, Appl1
8	3.2	89	4	US-09-149-476-451	Sequence 451, App
9	3.2	187	4	US-09-149-476-417	Sequence 417, App
10	3.2	222	1	US-07-869-933-12	Sequence 12, Appl
11	3.2	222	1	US-07-869-933-28	Sequence 28, Appl
12	3.2	222	4	US-09-103-663-12	Sequence 12, Appl

13	7	3.2	222	4	US-09-103-663-28	Sequence 28, Appl
14	7	3.2	267	2	US-08-557-128-4	Sequence 4, Appl1
15	7	3.2	267	2	US-09-242-690A-36	Sequence 36, Appl
16	7	3.2	326	1	US-07-603-133B-25	Sequence 25, Appl
17	7	3.2	380	4	US-08-857-076-110	Sequence 110, App
18	7	3.2	420	5	PCT-US96-10602-10	Sequence 10, Appl
19	7	3.2	427	4	US-09-134-001C-5143	Sequence 5143, Ap
20	7	3.2	554	4	US-09-321-276-4	Sequence 4, Appl1
21	7	3.2	554	4	US-08-916-481-2	Sequence 2, Appl1
22	7	3.2	563	4	US-08-916-481-3	Sequence 3, Appl1
23	7	3.2	663	4	US-09-106-293-3	Sequence 3, Appl1
24	7	3.2	663	4	US-08-209-603E-5	Sequence 5, Appl1
25	7	3.2	663	4	US-08-235-836C-70	Sequence 70, Appl1
26	7	3.2	663	4	US-08-235-836C-78	Sequence 78, Appl1
27	7	3.2	681	4	US-08-760-615-4	Sequence 4, Appl1
28	7	3.2	681	4	US-08-760-615-6	Sequence 6, Appl1
29	7	3.2	700	4	US-08-235-836C-74	Sequence 74, Appl1
30	7	3.2	1698	4	US-09-315-793-12	Sequence 12, Appl1
31	7	3.2	1724	4	US-08-857-076-12	Sequence 12, Appl1
32	6	2.8	11	1	US-08-338-634-4	Sequence 4, Appl1
33	6	2.8	11	1	US-08-787-547-67	Sequence 67, Appl
34	6	2.8	11	5	PCT-US95-16415-1	Sequence 1, Appl1
35	6	2.8	14	1	US-08-232-453A-40	Sequence 40, Appl1
36	6	2.8	15	4	US-09-027-900-3	Sequence 3, Appl1
37	6	2.8	17	2	US-08-497-599-2	Sequence 2, Appl1
38	6	2.8	17	2	US-08-497-599-18	Sequence 18, Appl
39	6	2.8	18	4	US-08-847-844A-50	Sequence 50, Appl
40	6	2.8	20	4	US-09-513-783A-146	Sequence 146, App
41	6	2.8	21	3	US-08-865-287-8	Sequence 8, Appl1
42	6	2.8	21	3	US-08-865-287-11	Sequence 11, Appl1
43	6	2.8	21	3	US-08-946-329A-40	Sequence 40, Appl1
44	6	2.8	22	1	US-08-241-853-37	Sequence 37, Appl
45	6	2.8	22	1	US-08-241-853-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
5182195-2

Patent No. 5182195

APPLICANT: NAKAHAMA, KAZUO; KAISHO, YOSHIMIRO; YOSHIMURA, KOJI

TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE

DEFICIENT YEASTS

NUMBER OF SEQUENCES: 71

CURRENT APPLICATION NUMBER: US-08/269, 140

FILING DATE: 09-NOV-1988

SEQ ID NO: 2:

LENGTH: 18

5182195-2

Alignment Scores:

Pred. No.: 123 Length: 18

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.21% Indels: 0

DB: 6 Gaps: 0

US-09-835-992A-19 (1-714) x 5182195-2 (1-18)

OY 493 CTGTTGCATTTGTTTCTTA 513

Db 6 LeuenaAlaLeucysphelen 12

RESULT 2

US-09-149-476-534

Sequence 534, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002P1

1	CURRENT	APPLICATION	NUMBER	US/09/149,476
2	CURRENT	FILING DATE	1998-09-08	
3	EARLIER	APPLICATION	NUMBER:	PCT/US98/04498
4	EARLIER	FILING DATE:	1998-03-06	
5	EARLIER	APPLICATION	NUMBER:	60/040,162
6	EARLIER	FILING DATE:	1997-03-07	
7	EARLIER	APPLICATION	NUMBER:	60/040,333
8	EARLIER	FILING DATE:	1997-03-07	
9	EARLIER	APPLICATION	NUMBER:	60/038,621
10	EARLIER	FILING DATE:	1997-03-07	
11	EARLIER	APPLICATION	NUMBER:	60/040,626
12	EARLIER	FILING DATE:	1997-03-07	
13	EARLIER	APPLICATION	NUMBER:	60/040,334
14	EARLIER	FILING DATE:	1997-03-07	
15	EARLIER	APPLICATION	NUMBER:	60/040,336
16	EARLIER	FILING DATE:	1997-03-07	
17	EARLIER	APPLICATION	NUMBER:	60/040,163
18	EARLIER	FILING DATE:	1997-03-07	
19	EARLIER	APPLICATION	NUMBER:	60/047,600
20	EARLIER	FILING DATE:	1997-05-23	
21	EARLIER	APPLICATION	NUMBER:	60/047,615
22	EARLIER	FILING DATE:	1997-05-23	
23	EARLIER	APPLICATION	NUMBER:	60/047,597
24	EARLIER	FILING DATE:	1997-05-23	
25	EARLIER	APPLICATION	NUMBER:	60/047,502
26	EARLIER	FILING DATE:	1997-05-23	
27	EARLIER	APPLICATION	NUMBER:	60/047,633
28	EARLIER	FILING DATE:	1997-05-23	
29	EARLIER	APPLICATION	NUMBER:	60/047,583
30	EARLIER	FILING DATE:	1997-05-23	
31	EARLIER	APPLICATION	NUMBER:	60/047,617
32	EARLIER	FILING DATE:	1997-05-23	
33	EARLIER	APPLICATION	NUMBER:	60/047,618
34	EARLIER	FILING DATE:	1997-05-23	
35	EARLIER	APPLICATION	NUMBER:	60/047,503
36	EARLIER	FILING DATE:	1997-05-23	
37	EARLIER	APPLICATION	NUMBER:	60/047,592
38	EARLIER	FILING DATE:	1997-05-23	
39	EARLIER	APPLICATION	NUMBER:	60/047,581
40	EARLIER	FILING DATE:	1997-05-23	
41	EARLIER	APPLICATION	NUMBER:	60/047,584
42	EARLIER	FILING DATE:	1997-05-23	
43	EARLIER	APPLICATION	NUMBER:	60/047,500
44	EARLIER	FILING DATE:	1997-05-23	
45	EARLIER	APPLICATION	NUMBER:	60/047,587
46	EARLIER	FILING DATE:	1997-05-23	
47	EARLIER	APPLICATION	NUMBER:	60/047,492
48	EARLIER	FILING DATE:	1997-05-23	
49	EARLIER	APPLICATION	NUMBER:	60/047,598
50	EARLIER	FILING DATE:	1997-05-23	
51	EARLIER	APPLICATION	NUMBER:	60/047,613
52	EARLIER	FILING DATE:	1997-05-23	
53	EARLIER	APPLICATION	NUMBER:	60/047,582
54	EARLIER	FILING DATE:	1997-05-23	
55	EARLIER	APPLICATION	NUMBER:	60/047,596
56	EARLIER	FILING DATE:	1997-05-23	
57	EARLIER	APPLICATION	NUMBER:	60/047,612
58	EARLIER	FILING DATE:	1997-05-23	
59	EARLIER	APPLICATION	NUMBER:	60/047,632
60	EARLIER	FILING DATE:	1997-05-23	
61	EARLIER	APPLICATION	NUMBER:	60/047,601
62	EARLIER	FILING DATE:	1997-05-23	
63	EARLIER	APPLICATION	NUMBER:	60/043,580
64	EARLIER	FILING DATE:	1997-04-11	
65	EARLIER	APPLICATION	NUMBER:	60/043,568
66	EARLIER	FILING DATE:	1997-04-11	
67	EARLIER	APPLICATION	NUMBER:	60/043,314
68	EARLIER	FILING DATE:	1997-04-11	
69	EARLIER	APPLICATION	NUMBER:	60/043,569
70	EARLIER	FILING DATE:	1997-04-11	
71	EARLIER	APPLICATION	NUMBER:	60/043,311
72	EARLIER	FILING DATE:	1997-04-11	
73	EARLIER	APPLICATION	NUMBER:	60/043,671

1	EARLIER	FILING DATE:	1997-04-11
2	EARLIER	APPLICATION NUMBER:	60/043, 67/4
3	EARLIER	FILING DATE:	1997-04-11
4	EARLIER	APPLICATION NUMBER:	60/043, 66/9
5	EARLIER	FILING DATE:	1997-04-11
6	EARLIER	APPLICATION NUMBER:	60/043, 312
7	EARLIER	FILING DATE:	1997-04-11
8	EARLIER	APPLICATION NUMBER:	60/043, 313
9	EARLIER	FILING DATE:	1997-04-11
10	EARLIER	APPLICATION NUMBER:	60/043, 67/2
11	EARLIER	FILING DATE:	1997-04-11
12	EARLIER	APPLICATION NUMBER:	60/043, 315
13	EARLIER	FILING DATE:	1997-04-11
14	EARLIER	APPLICATION NUMBER:	60/048, 97/4
15	EARLIER	FILING DATE:	1997-06-06
16	EARLIER	APPLICATION NUMBER:	60/056, 886
17	EARLIER	FILING DATE:	1997-08-22
18	EARLIER	APPLICATION NUMBER:	60/056, 877
19	EARLIER	FILING DATE:	1997-08-22
20	EARLIER	APPLICATION NUMBER:	60/056, 630
21	EARLIER	FILING DATE:	1997-08-22
22	EARLIER	APPLICATION NUMBER:	60/056, 87/8
23	EARLIER	FILING DATE:	1997-08-22
24	EARLIER	APPLICATION NUMBER:	60/056, 662
25	EARLIER	FILING DATE:	1997-08-22
26	EARLIER	APPLICATION NUMBER:	60/056, 87/2
27	EARLIER	FILING DATE:	1997-08-22
28	EARLIER	APPLICATION NUMBER:	60/056, 882
29	EARLIER	FILING DATE:	1997-08-22
30	EARLIER	APPLICATION NUMBER:	60/056, 637/3
31	EARLIER	FILING DATE:	1997-08-22
32	EARLIER	APPLICATION NUMBER:	60/056, 87/9
33	EARLIER	FILING DATE:	1997-08-22
34	EARLIER	APPLICATION NUMBER:	60/056, 880
35	EARLIER	FILING DATE:	1997-08-22
36	EARLIER	APPLICATION NUMBER:	60/056, 894
37	EARLIER	FILING DATE:	1997-08-22
38	EARLIER	APPLICATION NUMBER:	60/056, 911
39	EARLIER	FILING DATE:	1997-08-22
40	EARLIER	APPLICATION NUMBER:	60/056, 636
41	EARLIER	FILING DATE:	1997-08-22
42	EARLIER	APPLICATION NUMBER:	60/056, 87/4
43	EARLIER	FILING DATE:	1997-08-22
44	EARLIER	APPLICATION NUMBER:	60/056, 910
45	EARLIER	FILING DATE:	1997-08-22
46	EARLIER	APPLICATION NUMBER:	60/056, 864
47	EARLIER	FILING DATE:	1997-08-22
48	EARLIER	APPLICATION NUMBER:	60/056, 631
49	EARLIER	FILING DATE:	1997-08-22
50	EARLIER	APPLICATION NUMBER:	60/056, 845
51	EARLIER	FILING DATE:	1997-08-22
52	EARLIER	APPLICATION NUMBER:	60/056, 892
53	EARLIER	FILING DATE:	1997-08-22
54	EARLIER	APPLICATION NUMBER:	60/047, 599
55	EARLIER	FILING DATE:	1997-05-23
56	EARLIER	APPLICATION NUMBER:	60/047, 588
57	EARLIER	FILING DATE:	1997-05-23
58	EARLIER	APPLICATION NUMBER:	60/047, 585
59	EARLIER	FILING DATE:	1997-05-23
60	EARLIER	APPLICATION NUMBER:	60/047, 586
61	EARLIER	FILING DATE:	1997-05-23

EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:
Pred. No.: 122 Length: 19
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-149-476-534 (1-19)

OY 624 TTACTGTCCTAATTTATTC 644
Db 4 LeuLeuPheLeuIleLeuPhe 10

RESULT 3

5182195-7

Patent No. 5182195

APPLICANT: NAKAHAMA, KAZUO; KAISHO, YOSHIHIRO; YOSHIMURA, KOJI

TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE

DEFICIENT YEASTS

NUMBER OF SEQUENCES: 71

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/269,140

FILING DATE: 09-NOV-1988

SEQ ID NO: 7:
LENGTH: 21
5182195-7

Alignment Scores:
Pred. No.: 120 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 6 Gaps: 0

US-09-835-992A-19 (1-714) x 5182195-7 (1-21)

OY 493 CTGTCGATGCTGTTCTTA 513
Db 6 LeuLeuAlaLeuCyPheLeu 12

RESULT 4

US-09-288-143-88

Sequence 88, Application US/09288143

Patent No. 6433139

GENERAL INFORMATION:

APPLICANT: Brewer et al.

TITLE OF INVENTION: 53 Human Secreted Proteins

FILE REFERENCE: P2018P1

CURRENT APPLICATION NUMBER: US/09/288,143

CURRENT FILING DATE: 1999-04-08

EARLIER APPLICATION NUMBER: PCT/US98/21142

EARLIER FILING DATE: 1998-10-08

EARLIER APPLICATION NUMBER: 60/061,463

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,529

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/071,498

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,527

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,536

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,532

EARLIER FILING DATE: 1997-10-09

NUMBER OF SEQ ID NOS: 219

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 88

LENGTH: 44

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (44)

OTHER INFORMATION: Xaa equals stop translation

US-09-288-143-88

Alignment Scores:
Pred. No.: 107 Length: 44
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-288-143-88 (1-44)

OY 621 ATTTACTGTCCTAATTTTA 641

Db 5 IleLeuPheLeuIleLeu 11

RESULT 5

US-08-488-961-6

Sequence 6, Application US/08488961

Patent No. 5606042

GENERAL INFORMATION:

APPLICANT: Smith, Daniel S.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Reising, Ehlington, Barnard & Perry
STREET: P.O. Box 4390
CITY: Troy
STATE: Michigan
COUNTRY: US
ZIP: 48099-4390
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,961
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-320 (UMC)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 689-3500
TELEFAX: (810) 689-4071
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-488-961-6

Alignment Scores:
Pred. No.: 101 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x US-08-488-961-6 (1-62)

QY 496 TTGCATGTGTTCTTACTT 516
DB 20 LeuAlaLeuCysPheLeuLeu 26

RESULT 6
US-08-973-297-6
Sequence 6, Application US/08973297
Patent No. 6184017
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Kohn & Associates
STREET: 30500 No. 6184017Western Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/973,297
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994.00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5055
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-973-297-6

Alignment Scores:
Pred. No.: 101 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-19 (1-714) x US-08-973-297-6 (1-62)

QY 496 TTGCATGTGTTCTTACTT 516
DB 20 LeuAlaLeuCysPheLeuLeu 26

RESULT 7
PCT-US96-06511-6
Sequence 6, Application PC/TUS9606511
GENERAL INFORMATION:
APPLICANT: Smith, Daniel S.
TITLE OF INVENTION: Glycine and Phaseolus
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Kohn & Associates
STREET: 30500 Northwestern Hwy., Suite 410
CITY: Farmington Hills
STATE: Michigan
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/06511
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: 0994.00050
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 539-5055
TELEFAX: (810) 539-5055
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
PCT-US96-06511-6

Alignment Scores:

```
Pred. No.: 101 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 5 Gaps: 0

US-09-835-992a-19 (1-714) x PCT-US96-06511-6 (1-62)
OY 496 TTGCATTCGTCTTCTACTT 516
|||||
Db 20 LeuAlaLeuCySpheLeuLeu 26

RESULT 8
US-09-149-476-451
: Sequence 451, Application US/09149476
: Patent No. 6420526
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: 186 Human Secreted proteins
: FILE REFERENCE: P2002P1
: CURRENT APPLICATION NUMBER: US/09/149,476
: EARLIER FILING DATE: 1998-09-08
: EARLIER APPLICATION NUMBER: PCT/US98/04493
: EARLIER FILING DATE: 1998-03-06
: EARLIER APPLICATION NUMBER: 60/040,162
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,333
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/038,621
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,626
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,334
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,336
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,163
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/047,600
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,615
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,597
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,502
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,633
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,583
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,617
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,618
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,503
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,592
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,581
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,584
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,500
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,587
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,492
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,598
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,613
: EARLIER FILING DATE: 1997-05-23

: EARLIER APPLICATION NUMBER: 60/047,582
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,596
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,612
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,632
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,601
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/043,580
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,568
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,314
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,569
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,311
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,671
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,674
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,669
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,312
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,313
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,672
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/043,315
: EARLIER FILING DATE: 1997-04-11
: EARLIER APPLICATION NUMBER: 60/048,974
: EARLIER FILING DATE: 1997-06-06
: EARLIER APPLICATION NUMBER: 60/056,886
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,877
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,889
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,893
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,630
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,878
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,662
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,872
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,882
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,637
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,903
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,888
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,879
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,880
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,894
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,911
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,636
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,874
: EARLIER FILING DATE: 1997-08-22
: EARLIER APPLICATION NUMBER: 60/056,910
```

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:
Pred. No.: 95.7 Length: 89
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-149-476-451 (1-89)

OY 537 CTTAAATGATCTTCAGCAG 557
Db 62 Leuylsleuasnleungin 68

RESULT 9
US-09-149-476-417
Sequence 417, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:

Pred. No.: 85
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%

Length: 187
Matches: 7
Conservative: 0
Mismatch: 0

Query Match: 3.23% Indels: 0
DB: 4 Gaps: 0
US-09-835-992a-19 (1-714) x US-09-149-476-417 (1-187)
QY 164 TTTGTTATATTTCTTTT 144
DB 6 phevalllellepheuphe 12
RESULT 10
US-07-869-933-12
; Sequence 12, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGIN: SOURCE:
; ORGANISM: Rat
; STRAIN: FCRI alpha subunit
US-07-869-933-12
Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-19 (1-714) x US-07-869-933-12 (1-222)
QY 83 AATAACATTAGTATTAGAAA 103
DB 134 AsnAsnIleSerIleArgLys 140
RESULT 11
US-07-869-933-28
; Sequence 28, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:

APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 222 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGIN: SOURCE:
; STRAIN: alpha subunit
US-07-869-933-28
Alignment Scores:
Pred. No.: 82.8 Length: 222
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-19 (1-714) x US-07-869-933-28 (1-222)
QY 83 AATAACATTAGTATTAGAAA 103
DB 134 AsnAsnIleSerIleArgLys 140
RESULT 12
US-09-103-663-12
; Sequence 12, Application US/09103663D
; Patent No. 6171803
; GENERAL INFORMATION:
; APPLICANT: Kinet et al.
; TITLE OF INVENTION: Isolation, characterization, and use of the human beta
; TITLE OF INVENTION: subunit of the high affinity receptor for
; FILE REFERENCE: 50490
; CURRENT APPLICATION NUMBER: US/09/103,663D
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 07/869,933
; EARLIER FILING DATE: 1992-04-16
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Rattus sp.

US-09-103-663-12

Alignment Scores:	
Pred. No.:	82.8
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.21%
DB:	4
Length:	222
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-835-992A-19 (1-714) x US-09-103-663-12 (1-222)

Oy	83	AATAACATTACTATTAGAAA	103
Db	134	AsnAsnIleSerIleArgLys	140

RESULT 13
115-09-103-663-28

```

? Sequence 28. Application US/09103663D
? Patent No. 6171803
?
? GENERAL INFORMATION:
?
? APPLICANT: Kinet et al.
?
? TITLE OF INVENTION: Isolation, characterization, and use of the human beta2
?
? TITLE OF INVENTION: subunit of the high affinity
?
? TITLE OF INVENTION: Immunoglobulin E
?
? FILE REFERENCE: 50490
?
? CURRENT APPLICATION NUMBER: US/09/103,663D
?
? CURRENT FILING DATE: 1998-06-23
?
? EARLIER APPLICATION NUMBER: 07/869,933
?
? EARLIER FILING DATE: 1992-04-16
?
? NUMBER OF SEQ ID NOS: 35
?
? SOFTWARE: PatentIn Ver. 2.1
?
? SEQ ID NO 28
?
? LENGTH: 222
?
? TYPE: PR1
?
? ORGANISM: Homo sapiens
?
? OS-09-103-663-28

```

Alignment Scores:	
Pred. No.:	82.8
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.21%
DB:	4
Length:	2222
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-835-992A-19 (1-714) x US-09-103-663-28 (1-222)

QY	83	AATAACATTACTATTAGAAA	103
Db	134	AsnAsnIleSerIleargLys	140

RESULT 14
HE-09-557-128-A

Sequence 4, Application US/08557128
Patent No. 5849524
GENERAL INFORMATION:
APPLICANT: KONDO, Keiji
APPLICANT: KAJIWARA, Susumu
APPLICANT: MISAWA, No. 5849524hiko
TITLE OF INVENTION: TRANSFORMATION SYSTEMS FOR THE YEAST
TITLE OF INVENTION: CANDIDA UTILIS AND THE EXPRESSION OF HETEROLOGOUS GENES
TITLE OF INVENTION: THURENTH
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO PCT/JP95/01005
 FILING DATE: 25-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 7-129287
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-285823
 FILING DATE: 26-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 6-135015
 FILING DATE: 25-MAY-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 49441/108
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 267 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 OS-08-557-128-4

Alignment Scores:	
Pred. No.:	80.4
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.23%
DB:	2
Length:	265
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-835-992A-19 (1-714) x US-08-557-128-4 (1-267)

Qy	521	GA	CT	AG	T	A	G	A	A	C	A	C	A	A	T	501	
Db	80	G	L	U	S	E	R	L	Y	S	L	S	H	I	S	A	86

RESULT 15
HS-09-242

```

? Sequence 36, Application US/09242690A
? Patent No. 6284534
?
? GENERAL INFORMATION:
?
? APPLICANT: KONDO, KEIJI
? APPLICANT: MIURA, YUTAKA
? TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
? TITLE OF INVENTION: SAME
? FILE REFERENCE: 049441/0118
? CURRENT APPLICATION NUMBER: US/09/242,690A
? CURRENT FILING DATE: 1999-02-23
? PRIOR APPLICATION NUMBER: PCT/JP97/02924
? PRIOR FILING DATE: 1997-08-22
? PRIOR APPLICATION NUMBER: JP 8/241062
? PRIOR FILING DATE: 1996-08-23
? NUMBER OF SEQ ID NOS: 66
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 36
? LENGTH: 267
?
? TYPE: PRF
?
? ORGANISM: Candida utilis
?
? US-09-242-690A-36

```

Alignment Scores:

Pred. No.:	80.4	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.23%	Indels:	0
DB:	4	Gaps:	0

US-09-835-992A-19 (1-714) x US-09-242-690A-36 (1-267)

QY 521 GAACTAAGTAAAGAACACAAT 501

Db 80 GluLeuSerIysLysHisAsn 86

Search completed: January 14, 2003, 17:42:08
Job time : 12.0625 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 35.1708 Seconds

(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992a-19

Perfect score: 218
Sequence: 1 cgcagagaaagattactta.....cacagctgttcgcaattca 714

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-O/caga2.1/USPTO.spool/US09835992/rumat.14012003.161525.3979/app_query.fasta.1.3932
-DB=A_Geneseq.101002 -OPMT=fastan -SUFFIX=olig.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THRESH=SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFM=pro -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USR=US09835992.ecgn.1.121.0.rumat.14012003.161525.3979 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-YCAPOP=60 -YCAPEXT=60 -DELop=6 -DELExt=7

Database: A_Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. *Score* Query Match Length DB ID Description

C	1	8	3.7	29	19	AAW40006	Peptide effecting
	2	8	3.7	159	22	AB831093	Peptide #3744 enco
	3	8	3.7	159	22	AB836288	Peptide #3794 enco
	4	8	3.7	159	22	AB821653	Protein #3652 enco
	5	8	3.7	159	22	AAW57060	Human brain expres
	6	8	3.7	159	22	AAW69450	Human bone marrow
	7	8	3.7	159	22	AAW17284	Peptide #3718 enco
	8	8	3.7	159	22	AAW29784	Peptide #3621 enco
	9	8	3.7	159	22	AAW04971	Peptide #3653 enco
	10	8	3.7	159	23	ABG39070	Human peptide enco
	11	8	3.7	342	13	AAW25835	Human platelet act
	12	8	3.7	342	20	AAW49550	Human platelet act
	13	8	3.7	342	22	AB856381	Human platelet act
	14	8	3.7	342	23	AB877904	Non-endogenous hum
	15	8	3.7	342	23	AB877918	Amino acid sequenc
	16	8	3.7	485	22	AAW81002	Amino acid sequenc
	17	8	3.7	631	21	AAW56847	Rat neuronal gluta
	18	8	3.7	743	21	AAW56846	Arabidopsis thalia
	19	8	3.7	743	21	AAW56845	Arabidopsis thalia
	20	8	3.7	2870	21	AAW55559	Caenorhabditis ele
	21	8	3.7	3178	21	AAW55556	Caenorhabditis ele
	22	7	3.2	15	23	ABG62591	Eubacterial DNA po
	23	7	3.2	18	10	AAW92063	Modified type of e
	24	7	3.2	18	11	AAW05416	Improved hen egg w
	25	7	3.2	19	19	AAW74938	Human secreted pro
	26	7	3.2	21	9	AAW80973	Sequence of modifi
	27	7	3.2	21	10	AAW92065	N-terminal region
	28	7	3.2	21	11	AAW05417	Signal peptide
	29	7	3.2	22	20	AAW00284	Human secreted pro
	30	7	3.2	38	21	AAW63074	Human secreted pro
	31	7	3.2	43	21	AAW57883	Arabidopsis thalia
	32	7	3.2	44	20	AAW14432	Human secreted pro
	33	7	3.2	46	22	AAW22521	Novel human colon
	34	7	3.2	46	22	AAW2476	Human digestive sy
	35	7	3.2	47	22	AAW02178	Human polypeptide
	36	7	3.2	48	22	ABW5983	Human testicular a
	37	7	3.2	48	22	AAW5286	Human reproductive
	38	7	3.2	49	21	AAW6158	Arabidopsis thalia
	39	7	3.2	52	21	AAW00502	Human secreted pro
	40	7	3.2	55	20	AAW19543	Amino acid sequenc
	41	7	3.2	55	23	ABW4242	Human ORF protein
	42	7	3.2	62	18	AAW10324	Pinco bean alpha-D
	43	7	3.2	63	21	AAW07605	A human interleukin
	44	7	3.2	63	21	AAW07606	A human interleukin
	45	7	3.2	85	22	ABW70952	Drosophila melanog

ALIGNMENTS

RESULT 1

AAW40006 standard; peptide: 29 AA.

18-JUN-1998 (first entry)

Peptide effecting G-protein-coupled receptor activity.

G-protein-coupled receptor: GPCR; transmembrane domain; oligomerization;
therapeutic composition; GPCR function; receptor monomeric form;
multimeric form; inhibition; GPCR-mediated process; GPCR binding;
treatment; disease; platelet activating factor.

Synthetic.
Homo sapiens.
WO9800538-A2.

[illegible]

```

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI: 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
XX useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes
XX
XX
XX Claim 27; SEQ ID NO 14061; 327bp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a peptide encoded by a single exon nucleic acid
XX probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 159 AA;
XX
XX
XX Alignment Scores:
XX Pred. NO.: 42.6 Length: 159
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.67% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-835-992A-19 (1-714) x ABB31093 (1-159)
XX
XX QY 493 CTGTTGGCATTGTGTTTCTTACTT 516
XX |||||||||||||||||||
XX db 120 LeuLeuAlaLeuCySpheLeuLeu 127
XX
XX RESULT 3
XX ABB36288
XX ID ABB36288 standard; Peptide; 159 AA.
XX AC ABB36288;
XX
XX DT 04-FEB-2002 (first entry)
XX
XX DE Peptide #3794 encoded by human foetal liver single exon probe.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX 09-AUG-2001
XX

```

XX 30-JAN-2001; 2001WO-US00669.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 28923; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 159 AA;
 Alignment Scores:
 Pred. No.: 42.6 Length: 159
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.67% Indels: 0
 DB: 22 Gaps: 0
 US-09-835-992a-19 (1-714) x ABB36288 (1-159)
 QY 493 CTGTGGCATTGTTCTTACTT 516
 DB 120 LeuLeuAlaLeuCySpheLeuLeu 127
 RESULT 4
 ABB21653
 ID ABB21653 standard; Protein; 159 AA.
 AC ABB21653;
 XX
 DT 23-JAN-2002 (first entry)
 DE Protein #3652 encoded by probe for measuring heart cell gene expression.
 XX
 DE Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX
 OS Homo sapiens.
 OS
 PN WO200157274-A2.
 PD
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 04-OCT-2000; 2000US-0236359.

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488899/53.
 DR
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 15; SEQ ID NO 23423; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 159 AA;
 Alignment Scores:
 Pred. No.: 42.6 Length: 159
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.67% Indels: 0
 DB: 22 Gaps: 0
 US-09-835-992a-19 (1-714) x ABB21653 (1-159)
 QY 493 CTGTGGCATTGTTCTTACTT 516
 DB 120 LeuLeuAlaLeuCySpheLeuLeu 127
 RESULT 5
 AAM57060
 ID AAM57060 standard; Protein; 159 AA.
 AC AAM57060;
 XX
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29165.
 XX
 DE Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 OS
 PN WO200157275-A2.
 PD
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.

```
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 29165; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 159 AA;
XX
XX Alignment Scores:
XX Pred. No.: 42.6 Length: 159
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.67% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-835-992a-19 (1-714) x AAM69450 (1-159)
XX
XX QY 493 CTGTGGCATTGTTCTTACTT 516
XX |||||||||||||||||||
XX Db 120 LeuLeuAlaLeuCySpHeLeuLeu 127
XX
XX RESULT 6
XX AAM69450
XX ID AAM69450 standard; Protein; 159 AA.
XX
XX AC AAM69450;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29756.
XX
XX KW Human: bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00668.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI: 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
```

```
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 29756; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX
XX Sequence 159 AA;
XX
XX Alignment Scores:
XX Pred. No.: 42.6 Length: 159
XX Score: 8.00 Matches: 8
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.67% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-835-992a-19 (1-714) x AAM69450 (1-159)
XX
XX QY 493 CTGTGGCATTGTTCTTACTT 516
XX |||||||||||||||||||
XX Db 120 LeuLeuAlaLeuCySpHeLeuLeu 127
XX
XX RESULT 7
XX AAM17284
XX ID AAM17284 standard; Protein; 159 AA.
XX
XX AC AAM17284;
XX
XX DT 12-OCT-2001 (first entry)
XX
XX DE Peptide #3718 encoded by probe for measuring cervical gene expression.
XX
XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157278-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00670.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI: 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX Claim 27; SEQ ID NO 22110; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENP: see AAI10068-AAI8459). The present sequence is a peptide encoded
XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs
XX can be used to produce a single exon microarray, which can be used for
XX measuring human gene expression in a sample derived from human cervical
XX epithelial cells. By measuring gene expression, the probes are therefore
```

CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992A-19 (1-714) x AAM17284 (1-159)

OY 493 CTGTTGGCATTGTTTCTTACTT 516

Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 8

AAM29784
ID AAM29784 standard; Protein; 159 AA.

XX
AC AAM29784;

XX
DT 17-OCT-2001 (first entry)

XX
DE Peptide #3821 encoded by probe for measuring placental gene expression.

XX
KM Probe: microarray; human; placenta; antenatal diagnosis;

XX
OS Homo sapiens.

XX
PN W0200157272-A2.

XX
PD 09-AUG-2001.

XX
PF 30-JAN-2001; 2001WO-US00663.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-488897/53.

XX
XX Claim 27; SEQ ID NO 30053; 654pp; English.

XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see A113315-A115746). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.

XX
SQ Sequence 159 AA;

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
------------	------	---------	-----

Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992A-19 (1-714) x AAM29784 (1-159)

OY 493 CTGTTGGCATTGTTTCTTACTT 516

Db 120 LeuLeuAlaLeuCysPheLeuLeu 127

RESULT 9

AAM04971
ID AAM04971 standard; Protein; 159 AA.

XX
AC AAM04971;

XX
DT 09-OCT-2001 (first entry)

XX
DE Peptide #3653 encoded by probe for measuring breast gene expression.

XX
KM Probe: human; breast disease; breast cancer; development disorder;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX
OS Homo sapiens.

XX
PN W0200157270-A2.

XX
PD 09-AUG-2001.

XX
PF 29-JAN-2001; 2001WO-US00661.

XX
PR 04-FEB-2000; 2000US-0180312.

XX
PR 26-MAY-2000; 2000US-0207456.

XX
PR 30-JUN-2000; 2000US-0608408.

XX
PR 03-AUG-2000; 2000US-0632366.

XX
PR 21-SEP-2000; 2000US-0234687.

XX
PR 27-SEP-2000; 2000US-0236359.

XX
PR 04-OCT-2000; 2000GB-0024263.

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX
DR WPI; 2001-476286/51.

XX
XX Claim 27; SEQ ID NO 13711; 322pp; English.

XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see A100010-A110067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.

XX
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 159 AA;

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
------------	------	---------	-----

XX
SQ Sequence 159 AA;

XX
SQ Sequence 159 AA;

Alignment Scores:

Pred. No.:	42.6	Length:	159
------------	------	---------	-----

Query Match: 3.67% Indels: 0
 DB: 22 Gaps: 0

US-09-835-992a-19 (1-714) x AAM04971 (1-159)

OY 493 CTGTTGCAATGTTCTTACTT 516
 DB 120 LeuLeuAlaLeuCySpHeuLeu 127

RESULT 10
 ABG39070
 ID ABG39070 standard; Peptide: 159 AA.
 XX
 AC ABG39070;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 28735.
 XX
 KW Human; single exon probe: asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-063366P.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236529P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI: 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 27; SEQ ID No 28735; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene.

CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from Wipo at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 159 AA;
 XX

Alignment Scores:
 Pred. No.: 42.6 Length: 159
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.67% Indels: 0
 DB: 23 Gaps: 0

US-09-835-992a-19 (1-714) x ABG39070 (1-159)

OY 493 CTGTTGCAATGTTCTTACTT 516
 DB 120 LeuLeuAlaLeuCySpHeuLeu 127

RESULT 11
 AAR25835
 ID AAR25835 standard; Protein: 342 AA.
 XX
 AC AAR25835;
 XX
 DT 21-JAN-1993 (first entry)
 XX
 DE Human platelet activating factor receptor.
 XX
 KW PAF; coagulation; blood clotting; clone pHPAF.
 XX
 XX Homo sapiens.
 OS
 PN WO9212244-A.
 XX
 PD 23-JUL-1992.
 XX
 PF 27-DEC-1991; 91WO-JP01779.
 XX
 PR 27-DEC-1990; 90JP-0407119.
 PR 27-DEC-1990; 90JP-0407945.
 XX
 PA (NIBS) JAPAN TOBACCO INC.
 XX
 PI Honda Z, Nakamura M, Shimizu T;
 XX
 DR WPI: 1992-268667/32.
 DR N-PSDB; AAO26960.
 XX
 PT DNA coding for human or guinea pig platelet activating factor
 PT receptor - for prodn. of recombinant PAF receptor for screening
 PT potential PAF agonists-antagonists

XX C1alm 3; Fig 5; 50pp; Japanese.
 PS
 CC The cDNA coding for human PAF receptor was isolated in a 1780 bp
 CC clone designated pPAF. The clone consisted of a 112bp 5'-UTR (see
 CC AA026961), a 1029bp ORF (AA026960) and a 639bp 3'-UTR (see AA026962).
 CC The cDNA can be used to transform a suitable host cell line (e.g. COS
 CC cells) for expression of the PAF receptor protein. The receptor
 CC itself may be useful in screening of potential PAF agonists/
 CC antagonists and in assays for PAF. See also AA026957-Q26959.
 XX
 SQ Sequence 342 AA;
 Alignment Scores:
 Pred. No.: 38.9 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 13 Gaps: 0
 US-09-835-992A-19 (1-714) x AA026935 (1-342)
 OY 117 TCTTTTCTAGTTTCTATA 94
 DB 194 SerpPheLeuValPheLeuIle 201
 RESULT 12
 AA049550
 ID AA049550 standard; Protein: 342 AA.
 AC AA049550;
 DT 13-JAN-2000 (first entry)
 XX
 DE Human platelet activating factor receptor protein sequence.
 XX
 KW Human; coding sequence polymorphism; vascular pathology gene;
 KW polymorphic site; phenotype correlation; forensic; paternity testing;
 KW medicine; genetic analysis; vascular disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9950454-A2.
 PD 07-OCT-1999.
 XX
 PF 26-MAR-1999; 99WO-US06473.
 XX
 PR 01-APR-1998; 98US-0054272.
 XX
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Lander ES, Daley GO, Cargill M, Ireland JS, Rozen SG;
 XX
 DR WPI: 1999-620066/53.
 DR N-PSDB: AA232159.
 XX
 PT Determination of polymorphisms in genes, especially those identifying
 PT predisposition to vascular disease
 XX
 PS Disclosure: Fig 3; 134pp; English.
 XX
 CC AA232159 to AA232194 represent reference alleles for specifically
 CC claimed nucleic acid sequences from the present invention which comprise
 CC polymorphic sites as given in a table in the specification, selected
 CC from 92 single nucleotide polymorphisms in which the nucleotide at the
 CC polymorphic site is different from a nucleotide at the same site in a
 CC reference allele. The nucleic acids, and primers and probes, are used to
 CC identify polymorphisms, which may predispose an individual to disease,
 CC especially a vascular disease. They can also be used in phenotype
 CC correlations, forensics, paternity testing, medicine or genetic
 CC analysis. AA049550 to AA049573 represent the proteins which correspond

CC to some of the reference alleles.
 XX
 SQ Sequence 342 AA;
 Alignment Scores:
 Pred. No.: 38.9 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 20 Gaps: 0
 US-09-835-992A-19 (1-714) x AA049550 (1-342)
 OY 117 TCTTTTCTAGTTTCTATA 94
 DB 194 SerpPheLeuValPheLeuIle 201
 RESULT 13
 ABB56381
 ID ABB56381 standard; Protein: 342 AA.
 AC ABB56381;
 DT 18-FEB-2002 (first entry)
 XX
 DE Non-endogenous human GPCR protein, SMO ID NO: 555.
 XX
 KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
 KW constitutively activated GPCR; agonist; disease.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN WO200177172-A2.
 PD 18-OCT-2001.
 XX
 PF 05-APR-2001; 2001WO-US11098.
 XX
 PR 07-APR-2000; 2000US-195747P.
 XX
 PA (AREN-) ARENA PHARM INC.
 XX
 PI Lehmann-Brulnsma K, Llaw CW, Lin I;
 XX
 DR WPI: 2001-648759/74.
 DR N-PSDB: AB198017.
 XX
 PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
 PT disease treatment, comprises contacting candidate compounds with
 PT versions of GPCRs -
 XX
 PS Claim 1; Page 356-357; 394pp; English.
 XX
 CC The invention relates to G protein-coupled receptors (GPCRs) for which
 CC the endogenous ligand has been identified. Non-endogenous
 CC constitutively activated versions of known GPCRs are used in the
 CC invention for the direct identification of candidate compounds as
 CC receptor agonists, inverse agonists or partial agonists. Such
 CC agonists are useful as therapeutic agents for diseases or disorders
 CC associated with GPCRs. The present sequence is a non-endogenous
 CC version of a known human GPCR.
 XX
 SQ Sequence 342 AA;
 Alignment Scores:
 Pred. No.: 38.9 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 22 Gaps: 0

US-09-835-992a-19 (1-714) x ABB56381 (1-342)

QY 117 TCTTTTTCCTAGTTTTCATATA 94
 |||||
 Db 194 SerPhePheValPheLeuIle 201

RESULT 14

ABB77904
 ID ABB77904 standard; Protein; 342 AA.

AC ABB77904;
 XX
 DT 07-OCT-2002 (first entry)

DE Amino acid sequence of platelet activating factor receptor.
 XX
 KW Human; platelet activating factor receptor; PTAFR; gene; isogene;
 KW chromosome 1; inflammatory disease; coronary disease; cancer;
 KW receptor.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Misc-difference 28 Location/Qualifiers
 FT /note= "when PS2 is T, this residue is Phe"
 FT Misc-difference 224
 FT /note= "when PS3 is A, this residue is Asp"
 FT Misc-difference 338
 FT /note= "when PS5 is G, this residue is Ser"
 XX
 PN WO200251859-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 05-NOV-2001; 2001WO-US47441.
 XX
 PR 03-NOV-2000; 2000US-245633P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 XX
 PI Chew A, Choi JY, Koshy B;
 XX
 DR WPI: 2002-566672/60.
 DR N-PSDB: ABL59299.
 XX
 PS New genetic variants comprising haplotypes of the human Platelet
 PT Activating Factor Receptor (PTAFR) gene, useful for treating or
 PT screening drugs for treating e.g. inflammatory diseases, coronary
 PT diseases or cancer -
 XX
 PS Claim 29; Fig 3; 59pp; English.
 XX
 CC The present sequence represents a human Platelet Activating Factor
 CC Receptor (PTAFR). The PTAFR gene comprises polymorphic sites referred to
 CC as PS1-5 to designate the order in which they are located in the gene.
 CC Six isogenes of the PTAFR gene exist. The PTAFR gene is located on
 CC chromosome 1, and contains 1 exon. Polymorphisms PS3 and PS5 have
 CC previously been identified. PS3 and PS5 occur in the coding region.
 CC The polynucleotide comprising polymorphisms in the PTAFR gene is
 CC useful in screening candidate drugs to treat diseases related to PTAFR
 CC activity, e.g. inflammatory diseases, coronary diseases or cancer. The
 CC PTAFR isogenes are especially useful for treating these diseases. The
 CC methods and haplotypes are useful in improving the efficiency of drug
 CC discovery and development processes, or for designing clinical trials
 CC of candidate drugs for treating the specific condition or disease
 CC described above.
 CC
 XX
 SQ Sequence 342 AA;

Alignment Scores:
 Pred. No.: 38.9 Length: 342
 Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 23 Gaps: 0

US-09-835-992a-19 (1-714) x ABB77904 (1-342)

QY 117 TCTTTTTCCTAGTTTTCATATA 94
 |||||
 Db 194 SerPhePheValPheLeuIle 201

RESULT 15

ABB77918
 ID ABB77918 standard; Protein; 342 AA.

AC ABB77918;
 XX
 DT 07-OCT-2002 (first entry)

DE Amino acid sequence of platelet activating factor receptor variant.
 XX
 KW Human; platelet activating factor receptor; PTAFR; gene; isogene;
 KW chromosome 1; inflammatory disease; coronary disease; cancer;
 KW receptor.
 XX
 OS Homo sapiens.
 XX
 PN WO200251859-A2.
 XX
 PD 04-JUL-2002.
 XX
 PF 05-NOV-2001; 2001WO-US47441.
 XX
 PR 03-NOV-2000; 2000US-245633P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 XX
 PI Chew A, Choi JY, Koshy B;
 XX
 DR WPI: 2002-566672/60.
 DR
 XX
 PS New genetic variants comprising haplotypes of the human Platelet
 PT Activating Factor Receptor (PTAFR) gene, useful for treating or
 PT screening drugs for treating e.g. inflammatory diseases, coronary
 PT diseases or cancer -
 XX
 PS Claim 32; Page -: 59pp; English.
 XX
 CC The present sequence represents a human Platelet Activating Factor
 CC Receptor (PTAFR) polymorphic variant. The PTAFR gene has polymorphic
 CC sites referred to as PS1-5 to designate the order in which they are
 CC located in the gene. Six isogenes of the PTAFR gene exist. The PTAFR
 CC gene is located on chromosome 1, and contains 1 exon. Polymorphisms PS3
 CC and PS5 have previously been identified. PS3 and PS5 occur in the coding
 CC region. The polynucleotide comprising polymorphisms in the PTAFR gene is
 CC useful in screening candidate drugs to treat diseases related to PTAFR
 CC activity, e.g. inflammatory diseases, coronary diseases or cancer. The
 CC PTAFR isogenes are especially useful for treating these diseases. The
 CC methods and haplotypes are useful in improving the efficiency of drug
 CC discovery and development processes, or for designing clinical trials
 CC of candidate drugs for treating the specific condition or disease
 CC described above.
 CC note: this sequence does not appear in the specification; it was created
 CC using information provided.
 CC
 XX
 SQ Sequence 342 AA;

Alignment Scores:
 Pred. No.: 38.9 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0

Wed Jan 15 08:40:57 2003

DB:	23	Gaps:	0
US-09-835-992a-19 (1-714)	x	ABB77918 (1-342)	
Oy	117	TCCTTTCTAGTTTCTAATA	94
Db	194	SerPheLeuValPheLeuIle	201

Search completed: January 14, 2003, 17:27:17
Job time : 38.1708 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 6.8357 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-19

Perfect score: 218
Sequence: 1 cgcgcagaaagattatltta.....cacagtggtgcacatca 714

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 1940157 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+ n2p.model -DEV-xlp
-O/cgnt2.1/USPTO.spool/US09835992/runit.14012003.161528.4118/app.query.fasta_1.3932
-DB-published.Applications_AA -QFMT-fasten -SUFFIX-olig.rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITs-bits -START=1 -END=1 -MATRIX-oligo
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09835992 -GCN 1.1_28 -runit.14012003.161528.4118
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAL -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Published Applications_AA:
1: /cgnt2.6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgnt2.6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgnt2.6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgnt2.6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
5: /cgnt2.6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
6: /cgnt2.6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgnt2.6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
8: /cgnt2.6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
9: /cgnt2.6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
10: /cgnt2.6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
11: /cgnt2.6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
12: /cgnt2.6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgnt2.6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgnt2.6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query (Score)	Match	Length	ID	Description
1	8	3.7	159	US-09-864-761-36951	Sequence 36951, A
2	8	3.7	327	US-10-084-206-3	Sequence 3, Appl
3	7	3.2	81	US-10-002-144A-243	Sequence 243, App
4	7	3.2	127	US-09-764-877-1466	Sequence 1466, App

5	7	3.2	197	US-09-780-717-47	Sequence 47, Appl
6	7	3.2	241	US-09-740-668A-26	Sequence 26, Appl
7	7	3.2	267	US-09-908-855-36	Sequence 36, Appl
8	7	3.2	334	US-09-815-242-13249	Sequence 13249, A
9	7	3.2	380	US-09-205-658-110	Sequence 110, App
10	7	3.2	380	US-09-844-353A-110	Sequence 110, App
11	7	3.2	420	US-09-812-862-10	Sequence 10, Appl
12	7	3.2	481	US-09-815-242-10888	Sequence 10888, A
13	7	3.2	554	US-09-800-396-4	Sequence 4, Appl
14	7	3.2	1724	US-09-205-658-112	Sequence 12, Appl
15	7	3.2	1724	US-09-844-353A-112	Sequence 12, Appl
16	6	2.8	9	US-09-834-765-444	Sequence 44, App
17	6	2.8	10	US-08-452-843A-13	Sequence 13, Appl
18	6	2.8	10	US-09-834-765-82	Sequence 82, Appl
19	6	2.8	10	US-09-834-765-611	Sequence 61, App
20	6	2.8	11	US-09-909-460-67	Sequence 67, Appl
21	6	2.8	17	US-09-949-196-4	Sequence 4, Appl
22	6	2.8	22	US-09-949-196-32	Sequence 32, Appl
23	6	2.8	22	US-09-949-196-36	Sequence 36, Appl
24	6	2.8	22	US-09-949-196-40	Sequence 40, Appl
25	6	2.8	25	US-10-001-879-113	Sequence 113, App
26	6	2.8	28	US-09-925-299-1023	Sequence 1023, App
27	6	2.8	28	US-09-929-818-104	Sequence 104, App
28	6	2.8	30	US-09-071-838-227	Sequence 227, App
29	6	2.8	31	US-09-864-761-38920	Sequence 38920, A
30	6	2.8	32	US-09-864-761-47089	Sequence 47089, A
31	6	2.8	34	US-09-864-761-34176	Sequence 34176, A
32	6	2.8	36	US-09-864-761-47949	Sequence 47949, A
33	6	2.8	36	US-09-864-761-48325	Sequence 48325, A
34	6	2.8	37	US-09-864-761-44137	Sequence 44137, A
35	6	2.8	38	US-09-864-761-46306	Sequence 46306, A
36	6	2.8	39	US-09-764-847-937	Sequence 937, App
37	6	2.8	40	US-09-864-761-41296	Sequence 41296, A
38	6	2.8	40	US-09-864-761-42251	Sequence 42251, A
39	6	2.8	41	US-09-864-761-46409	Sequence 46409, A
40	6	2.8	43	US-09-864-761-44190	Sequence 44190, A
41	6	2.8	43	US-09-764-847-866	Sequence 866, App
42	6	2.8	44	US-09-864-761-47765	Sequence 47765, A
43	6	2.8	45	US-09-764-877-1815	Sequence 1815, App
44	6	2.8	46	US-09-764-878-157	Sequence 157, App
45	6	2.8	47	US-09-925-299-785	Sequence 785, App

ALIGNMENTS

RESULT 1
US-09-864-761-36951
Sequence 36951, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36951
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007374.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AAA85663.1, EVALUATE 1.00e-67
US-09-864-761-36951

Alignment Scores:
Pred. No.: 5.88 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-19 (1-714) x US-09-864-761-36951 (1-159)
QY 493 CTGTGCATGTCTTCTACTT 516
Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 2
US-10-084-206-3
; Sequence 3, Application US/10084206
; Patent No. US20020106741A1
; GENERAL INFORMATION:
; APPLICANT: Li et al.
; TITLE OF INVENTION: G-Protein Receptor HTNAD29
; FILE REFERENCE: P191D1C1
; CURRENT APPLICATION NUMBER: US/10/084,206
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US95/07288
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/468,534
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 09/399,095
; PRIOR FILING DATE: 1999-09-20
```

```

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 3
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-084-206-3

Alignment Scores:
Pred. No.: 5.21 Length: 327
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-19 (1-714) x US-10-084-206-3 (1-327)
QY 117 TCTTTTTCCTAGTTTCTAATA 94
Db 190 SerPheLeuValPheLeuLeu 197

RESULT 3
US-10-002-344A-243
; Sequence 243, Application US/10002344A
; Patent No. US20020172959A1
; GENERAL INFORMATION:
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Chen, Sei-yu
; APPLICANT: Liu, Chenghua
; APPLICANT: Turner, Leah
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
; FILE REFERENCE: DEX-0241
; CURRENT APPLICATION NUMBER: US/10/002,344A
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/242,998
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 243
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-002-344A-243

Alignment Scores:
Pred. No.: 61.2 Length: 81
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 9 Gaps: 0

US-09-835-992A-19 (1-714) x US-10-002-344A-243 (1-81)
QY 621 ATTTTACTGTTCTAATTTTA 641
Db 73 IleLeuLeuPheLeuLeuLeu 79

RESULT 4
US-09-764-877-1466
; Sequence 1466, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
```

SEQ ID NO 1466
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1466

Alignment Scores:
Pred. No.: 56.7 Length: 127
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-764-877-1466 (1-127)

QY 326 GCCTGCTCAAGCTCTGATC 346
|||||
Db 80 AlacysvalGlnAlaLeuIle 86

RESULT 5
US-09-780-717-47
Sequence 47, Application US/09780717
Patent No. US20010044941A1
GENERAL INFORMATION:
APPLICANT: Helentjaris, Tim
APPLICANT: Bates, Nic
TITLE OF INVENTION: of use
FILE REFERENCE: 035718/208677
CURRENT APPLICATION NUMBER: US/09/780,717
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,509
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 197
TYPE: PRT
ORGANISM: Glycine max
US-09-780-717-47

Alignment Scores:
Pred. No.: 52.6 Length: 197
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-780-717-47 (1-197)

QY 43 CAAGCATTTATTTATCTAT 63
|||||
Db 45 GlnserIleIleuSerTyf 51

RESULT 6
US-09-740-668a-26
Sequence 26, Application US/09740668A
Patent No. US20020076700A1
GENERAL INFORMATION:
APPLICANT: Shumkeis, Richard
TITLE OF INVENTION: No. US20020076700A1 polypeptides and nucleic acids encoding sam
FILE REFERENCE: 15966-537 CIP
CURRENT APPLICATION NUMBER: US/09/740,668A
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: PCT/US99/29584
PRIOR FILING DATE: 1999-12-17

PRIOR APPLICATION NUMBER: 09/465,512
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/113,485
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/112,837
PRIOR FILING DATE: 1998-12-18
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 241
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)-(241)
OTHER INFORMATION: wherein Xaa is any amino acid as defined in the specification.
US-09-740-668a-26

Alignment Scores:
Pred. No.: 50.9 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-740-668a-26 (1-241)

QY 338 GCCTGATCATATTTCTTTTA 358
|||||
Db 126 AlaIleuIleIlePheLeu 132

RESULT 7
US-09-908-855-36
Sequence 36, Application US/09908855
Patent No. US20020115220A1
GENERAL INFORMATION:
APPLICANT: KONDO, KEIJI
APPLICANT: MIURA, YUTARA
TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
FILE REFERENCE: 049441/0118
CURRENT APPLICATION NUMBER: US/09/908,855
CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: 09/242,690
PRIOR FILING DATE: 1999-02-23
PRIOR APPLICATION NUMBER: JP 8/241062
PRIOR FILING DATE: 1996-08-23
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 267
TYPE: PRT
ORGANISM: Candida utilis
US-09-908-855-36

Alignment Scores:
Pred. No.: 50 Length: 267
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-908-855-36 (1-267)

QY 521 GAAGTAAGTAAGAACACAT 501
|||||
Db 80 GluSerIleTySHisAsn 86

RESULT 8
US-09-815-242-13249
Sequence 13249, Application US/09815242

Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA-011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13249
LENGTH: 334
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13249
Alignment Scores:
Pred. No.: 48.1 Length: 334
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0
US-09-835-992A-19 (1-714) x US-09-815-242-13249 (1-334)
QY 402 GCATTCATGCTATCGAAGC 382
DB 121 AlAphehIsAlIleAtgAsn 127
RESULT 9
US-09-205-658-110
Sequence 110, Application US/09205658
Patent No. US20010029617A1
GENERAL INFORMATION:
APPLICANT: Ruukun, Gary
APPLICANT: Ogy, Scott
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351004
CURRENT APPLICATION NUMBER: US/09/205,658
CURRENT FILING DATE: 1998-12-03
EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15
EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07
EARLIER APPLICATION NUMBER: US98/10080
EARLIER FILING DATE: 1998-05-15
NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 110
LENGTH: 380

TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-205-658-110
Alignment Scores:
Pred. No.: 47.1 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0
US-09-835-992A-19 (1-714) x US-09-205-658-110 (1-380)
QY 108 CTAGTTTCTAATGACTAATG 88
DB 4 LeuValPheLeuIleLeuMet 10
RESULT 10
US-09-844-353A-110
Sequence 110, Application US/09844353A
Patent No. US20020037585A1
GENERAL INFORMATION:
APPLICANT: Ruukun, Gary
APPLICANT: Kimura, Koutaro
APPLICANT: Paterson, Garth
APPLICANT: Ogy, Scott
APPLICANT: Paradis, Suzanne
APPLICANT: Tissenbaum, Heidi
APPLICANT: Morris, Jason
APPLICANT: Kowsek, Allison
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
FILE REFERENCE: 00786/351005
CURRENT APPLICATION NUMBER: US/09/844,353A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 08/857,076
PRIOR FILING DATE: 1997-05-15
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 110
LENGTH: 380
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-844-353A-110
Alignment Scores:
Pred. No.: 47.1 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0
US-09-835-992A-19 (1-714) x US-09-844-353A-110 (1-380)
QY 108 CTAGTTTCTAATGACTAATG 88
DB 4 LeuValPheLeuIleLeuMet 10
RESULT 11
US-09-812-862-10
Sequence 10, Application US/09812862
Patent No. US20020035081A1
GENERAL INFORMATION:
APPLICANT: Wands, Jack R.
APPLICANT: Scaglioni, Pier Paolo
APPLICANT: Melegari, Margherita
TITLE OF INVENTION: INHIBITION OF HEPATITIS B REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/812,862
FILING DATE: 20-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,073
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/492,489
FILING DATE: 20-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/282001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-812-862-10

Alignment Scores:
Pred. No.: 46.3 Length: 420
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-812-862-10 (1-420)

OY 514 CTTAGTCTCCGCAAGCAAA 534
|||||
Db 292 LeuserSerProlysglyLys 298

RESULT 12
US-09-815-242-10888
Sequence 10888, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10888
LENGTH: 481
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10888

Alignment Scores:
Pred. No.: 45.3 Length: 481
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-815-242-10888 (1-481)

OY 205 AAATGCAATCTGAACCAA 225
|||||
Db 399 LysThrAlaSerGluSerLys 405

RESULT 13
US-09-800-396-4
Sequence 4, Application US/09800396
Patent No. US20020065395A1
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola
TITLE OF INVENTION: NOVEL RESPONSE REGULATOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/800,396
FILING DATE: 06-Mar-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/879,531
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER: GM10018
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-800-396-4

Alignment Scores:
Pred. No.: 44.2 Length: 554
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-800-396-4 (1-554)

OY 111 TTCTAGTTTCTACTATG 91
Db 6 PheValPheUleUleUmet 12

RESULT 14
US-09-205-658-12
; Sequence 12, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-205-658-12

Alignment Scores:
Pred. No.: 36.5 Length: 1724
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-205-658-12 (1-1724)

OY 108 CTAGTTTCTACTACTATG 88
Db 1068 LeuValPheUleUleUmet 1074

RESULT 15
US-09-844-353A-12
; Sequence 12, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutaro
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351005
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1724
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-844-353A-12

Alignment Scores:
Pred. No.: 36.5 Length: 1724
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-19 (1-714) x US-09-844-353A-12 (1-1724)

OY 108 CTAGTTTCTACTACTATG 88
Db 1068 LeuValPheUleUleUmet 1074

Search completed: January 14, 2003, 18:08:15
Job time: 10.8357 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 23.925 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992a-19

Perfect score: 218
Sequence: 1 cgcgcgaagaaatatttca.....cacagtggttcacattca 714

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=xlp
-O/cgr2.1/USPTO.spool/US09835992/rnatc_14012003.161526.4038/app_query.fasta.1.3932
-DB=PIR_73 -QMT=fastan -SUFFIX=olig.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptlo
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US09835992.qcgn.1.1.192.qrunatc_14012003.161526.4038 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database:

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	8	3.7	195	2	E71000
C 2	8	3.7	342	2	A40191
C 3	8	3.7	343	2	T19082
C 4	8	3.7	393	2	H64233
C 5	8	3.7	512	2	E83060
C 6	8	3.7	791	2	H72552
C 7	8	3.7	1118	1	A49724
C 8	8	3.7	1183	2	F90559
C 9	8	3.7	2329	2	T28125
C 10	7	3.2	86	2	S34538
C 11	7	3.2	85	2	S21595
C 12	7	3.2	105	2	E84075
C 13	7	3.2	113	2	C70342
C 14	7	3.2	120	2	T30769

15	7	3.2	126	2	E72150	BCL protein - vari
16	7	3.2	164	1	MMMU1	2S albumin 1 precu
17	7	3.2	164	1	MMMU3	2S albumin 3 precu
C 18	7	3.2	194	2	C69933	negative regulatio
C 19	7	3.2	201	2	G90674	probable transcrip
C 20	7	3.2	201	2	B85525	probable transcrip
C 21	7	3.2	201	2	F64758	yhnd protein - Esc
C 22	7	3.2	205	2	T08398	calmodulin homolog
C 23	7	3.2	207	2	T24267	hypothetical prote
C 24	7	3.2	208	1	A69313	conserved hypotet
C 25	7	3.2	214	2	G70348	conserved hypotet
C 26	7	3.2	236	1	B64212	hypothetical prote
C 27	7	3.2	237	2	C81291	probable integral
C 28	7	3.2	245	2	A30154	IGF receptor alpha
C 29	7	3.2	249	2	C81449	probable molybde
C 30	7	3.2	254	2	D69140	hypothetical prote
C 31	7	3.2	271	2	F72379	sugar ABC transpor
C 32	7	3.2	278	2	B75442	3-hydroxybutyryl-C
C 33	7	3.2	279	2	T20443	hypothetical prote
C 34	7	3.2	290	2	G72203	sugar ABC transpor
C 35	7	3.2	307	2	G64350	quinolinate synthe
C 36	7	3.2	314	2	T45010	probable site-spec
C 37	7	3.2	314	2	S73717	hypothetical prote
C 38	7	3.2	314	2	E84970	pseudouridylylate sy
C 39	7	3.2	317	2	B71110	hypothetical prote
C 40	7	3.2	319	2	D90589	hypothetical prote
C 41	7	3.2	320	2	T30894	lipase homolog T0
C 42	7	3.2	326	1	VGXRRH	glycoprotein VP7 P
C 43	7	3.2	326	1	VGXRRR	glycoprotein VP7 P
C 44	7	3.2	326	1	A44891	glycoprotein VP7 P
C 45	7	3.2	334	2	B98019	conserved hypotet

ALIGNMENTS

RESULT 1

E71000
hypothetical protein PH1301 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 18-Aug-2000
C:Accession: E71000
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haino, Y.; Yamamoto, S.; Se
M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71000
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-195 <RAW>
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by Genba
C:Genetics:
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1301

Alignment Scores:

Prod. No.: 10.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.69%
DB: 2
Gaps: 0

US-09-835-992a-19 (1-714) x E71000 (1-195)

Qy 120 TTATCTTTTTCCTAGTTTCTA 97

Db 99 LeuserPheleValPheleu 106

RESULT 2

A40191

platelet-activating factor receptor - human
 C:Species: Homo sapiens (man)
 C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 20-Jun-2000
 C:Accession: A40191; JH0479; A41079; J01359; A42831; I51923
 R:Kunz, D.; Gerard, N.P.; Gerard, C.
 J. Biol. Chem. 267, 9101-9106, 1992
 A>Title: The human leukocyte platelet-activating factor receptor. cDNA cloning, cell surface expression, and functional characterization
 A:Reference number: A40191; MUID:92250505; PMID:1374385
 A:Accession: A40191
 A:Molecule type: mRNA
 A:Residues: 1-342 <KUN>
 A:Cross-references: GB:M76674; NID:9456293; PIDN:AAA60002.1; PID:9456294
 R:Ye, R.D.; Prossnitz, E.R.; Zou, A.; Cochane, C.G.
 Biochem. Biophys. Res. Commun. 180, 105-111, 1991
 A>Title: Characterization of a human cDNA that encodes a functional receptor for platelet-activating factor
 A:Reference number: JH0479; MUID:92028922; PMID:1656963
 A:Accession: JH0479
 A:Molecule type: mRNA
 A:Residues: 1-342 <YER>
 A:Cross-references: GB:M80436; NID:9189537; PIDN:AAA60001.1; PID:9189538
 A:Experimental source: granulocyte, cell line HL-60 all
 R:Nakamura, M.; Honda, Z.; Izumi, T.; Sakanaka, C.; Mutoh, H.; Minami, M.; Bito, H.; Sey, J. Biol. Chem. 266, 20400-20405, 1991
 A>Title: Molecular cloning and expression of platelet-activating factor receptor from human platelets
 A:Reference number: A41079; MUID:92041873; PMID:1657923
 A:Accession: A41079
 A:Molecule type: not compared with conceptual translation
 A:Status: not compared with conceptual translation
 A:Residues: 1-342 <NAK>
 A:Cross-references: GB:D10202; GB:D90433; NID:9219975; PIDN:BA01050.1; PID:9219976
 R:Sugimoto, T.; Tsuchimochi, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurauchi, Y. Biochem. Biophys. Res. Commun. 189, 617-624, 1992
 A>Title: Molecular cloning and characterization of the platelet-activating factor receptor
 A:Reference number: J01359; MUID:93112021; PMID:1281995
 A:Accession: J01359
 A:Molecule type: mRNA
 A:Residues: 1-315, 'N', 317-342 <SUG>
 A:Experimental source: heart
 A>Note: The authors translated the codon AAT for residue 316 as Lys
 R:Seyfried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W. Genomics 13, 832-834, 1992
 A>Title: The human platelet-activating factor receptor gene (P1AFR) contains no introns
 A:Reference number: A42831; MUID:92347866; PMID:1322356
 A:Accession: A42831
 A:Molecule type: DNA
 A:Residues: 1-226, 'TG', 229-342 <SEY>
 A:Cross-references: GB:M8177; NID:9190697; PIDN:AAA60214.1; PID:9190698
 A>Note: Sequence extracted from NCBI backbone (NCBIN:109813, NCBI:109814)
 R:Chase, P.B.; Halonen, M.; Regan, J.W. Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
 A>Title: Cloning of a human platelet-activating factor receptor gene: evidence for an intron
 A:Reference number: I51923; MUID:93192035; PMID:8383507
 A:Accession: I51923
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-342 <RES>
 A:Cross-references: GB:S56396; NID:9298580; PIDN:AA825755.1; PID:9298581
 C:Genetics:
 A:Gene: GDB:P7AFR
 A:Cross-references: GDB:128806; OMIM:173393
 A:Map position: 1p35-1p34.3
 C:Superfamily: ATP receptor P2U
 C:Keywords: G protein-coupled receptor; transmembrane protein
 F:17-38/Domain: transmembrane #status predicted <TRI>
 F:54-75/Domain: transmembrane #status predicted <TRI>
 F:92-113/Domain: transmembrane #status predicted <TRI>
 F:134-155/Domain: transmembrane #status predicted <TRV>
 F:184-205/Domain: transmembrane #status predicted <TRV>
 F:233-253/Domain: transmembrane #status predicted <TVI>
 F:277-297/Domain: transmembrane #status predicted <TVI>
 Alignment Scores:
 Pred. No.: 10 Length: 342

Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 2 Gaps: 0
 US-09-835-992a-19 (1-714) x A40191 (1-342)
 QY 117 TCTTTTCTAGTTTTCATATA 94
 DB 194 SerPhePheValPheLeuLeu 201
 RESULT 3
 T19082
 hypothetical protein C08B6.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19082
 R:Wilkinson, J. submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19070
 A:Accession: T19082
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-343 <WIL>
 A:Cross-references: EMBL:Z72502; PIDN:CAA96591.1; GSPDB:GN00023; CESP:C08B6.9
 A:Experimental source: clone C08B6
 C:Genetics:
 A:Gene: CESP:C08B6.9
 A:Map position: 5
 A:Introns: 25/2; 84/3; 174/2; 290/2
 Alignment Scores:
 Pred. No.: 10 Length: 343
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 2 Gaps: 0
 US-09-835-992a-19 (1-714) x T19082 (1-343)
 QY 646 CAGATTAATAATGAGACGTATA 623
 DB 23 GlnAsnLysIleArgAsnSerLys 30
 RESULT 4
 H64233
 hypothetical protein MG306 - Mycoplasma genitalium
 C:Species: Mycoplasma genitalium
 C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
 C:Accession: H64233
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.M.; Fuhmann, J.; Nguyen, D.; Uitterlacker, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C. Science 270, 397-403, 1995
 A>Title: The minimal gene complement of Mycoplasma genitalium.
 A:Reference number: A64200; MUID:96026346; PMID:7569993
 A:Accession: H64233
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-393 <TIGR>
 A:Cross-references: GB:U39712; GB:L43967; NID:g1046005; PID:g1046006; TIGR:MG306
 A:Experimental source: strain G-37
 C:Genetics:
 A:Genetic code: SGC3
 Alignment Scores:
 Pred. No.: 9.88 Length: 393
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0

```
DB: 2 Gaps: 0
US-09-835-992a-19 (1-714) x H64233 (1-393)
OY 117 TCTTTTCTACTTTTCTAATA 94
Db 19 SerpheeLeuValpheeLeu 26
RESULT 5
E83060
Iron (III)-transport system permease HltB PA4688 [imported] - Pseudomonas aeruginosa (str
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83060
R:Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,
; Loy, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83060
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE004882; GB:AE004091; NID:99950939; PIDN:BA08075.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: hltB; PA4688
C:Superfamily: sfub protein
Alignment Scores:
Pred. No.: 9.61 Length: 512
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
Db: 2 Gaps: 0
US-09-835-992a-19 (1-714) x E83060 (1-512)
OY 493 CTCTTGGCATTCGTTCTTACTT 516
Db 230 LeuValLeuValpheeLeu 237
RESULT 6
H72552
hypothetical protein APE1708 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: H72552
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Taket
awa, H.; Takemura, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72552
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-791 <KAW>
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BA080709.1; PID:dl044495; PID:9510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1708
Alignment Scores:
Pred. No.: 9.19 Length: 791
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
Db: 2 Gaps: 0
US-09-835-992a-19 (1-714) x H72552 (1-791)
OY 90 TTAGTATTAGAAAACTAGGAAA 113
Db 134 LeuValLeuGluValLeuGluGly 141
RESULT 7
A49724
protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type II precursor - human
N:Alternate names: protein-tyrosine-phosphatase, stomach cancer-associated type 1; SA
C:Species: Homo sapiens (man)
C:Date: 03-May-1994 #sequence_revision 23-Feb-1996 #text_change 16-Jun-2000
C:Accession: A49724
R:Matocak, T.; Suzuki, T.; Uchida, T.; Inazawa, J.; Ariyama, T.; Matsuda, K.; Horita
J. Biol. Chem. 269, 2075-2081, 1994
A:Title: Molecular cloning of a human transmembrane-type protein tyrosine phosphatase
A:Reference number: A49724; MUID:94124561; PMID:8294459
A:Accession: A49724
A:Molecule type: mRNA
A:Residues: 1-1118 <MATO>
A:Cross-references: GB:D15049; NID:9475003; PIDN:BA03645.1; PID:9475004
C:Genetics:
A:Gene: GDB:PPP8; SAP-1
A:Cross-references: GDB:305504
A:Map position: 19q13.4-19q13.4
A:Note: highly expressed in colon and pancreatic cancer cells but not in the normal c
C:Superfamily: protein-tyrosine-phosphatase, receptor type II; fibronectin type III re
C:Keywords: carcinogenesis; duplication; glycoprotein; phosphoprotein; phosphoric mon
F:1-27/Domain: signal sequence #status predicted <SIG>
F:27-110/Domain: fibronectin type III repeat homology <3FNA>
F:28-118/Domain: fibronectin type III repeat homology <3FNB>
F:28-761/Domain: extracellular #status predicted <EXT>
F:116-199/Domain: fibronectin type III repeat homology <3FNC>
F:205-289/Domain: fibronectin type III repeat homology <3FND>
F:296-379/Domain: fibronectin type III repeat homology <3FNE>
F:385-468/Domain: fibronectin type III repeat homology <3FNF>
F:474-558/Domain: fibronectin type III repeat homology <3FNG>
F:564-658/Domain: fibronectin type III repeat homology <3FNG>
F:667-737/Domain: fibronectin type III repeat homology <3FNG>
F:762-778/Domain: transmembrane #status predicted <TMN>
F:779-1118/Domain: intracellular #status predicted <INT>
F:846-1070/Domain: protein-tyrosine-phosphatase homology <PPPI>
F:35,78,83,107,132,149,172,196,203,286,304,312,329,352,376,383,401,436,439,470,490,55
F:1022/Active site: Cys (phosphocysteine intermediate) #status predicted
F:1028/Binding site: substrate phosphate (Arg) #status predicted
Alignment Scores:
Pred. No.: 8.86 Length: 1118
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
Db: 1 Gaps: 0
US-09-835-992a-19 (1-714) x A49724 (1-1118)
OY 618 GGCATTTACTGTTCTAATTTA 641
Db 763 GlyLeuLeuValpheeLeu 770
RESULT 8
F90559
conserved hypothetical protein MYPU_3820 [imported] - Mycoplasma pulmonis (strain UAB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: F90559
R:Chamblaud, I.; Helliou, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: F90559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1183 <KUR>
```

A:Cross-references: GB:AL45566; PID:914089796; PIDN:CAC13555.1; GSPDB:GN00153
A:Experimental source: strain UAB CT1P
C:Genetics:
A:Gene: MYD_3820
A:Genetic code: SGC3

Alignment Scores:
Pred. No.: 8.81 Length: 1183
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-19 (1-714) x P90559 (1-1183)

OY 208 TTTTCAATTATATCTCATTTCA 185
|||||
DB 577 PheSerAsnTyrAsnLeuIleSer 584

RESULT 9
T28125
hypothetical protein ZK945.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28125
R:Wilkinson, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z20472
A:Accession: T28125
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-2329 <N1L>
A:Cross-references: EMBL:Z48544; NID:9695493; PIDN:CAA8442.1; GSPDB:GN00020; CESP:ZK945
A:Experimental source: clone ZK945
C:Genetics:
A:Gene: CESP:ZK945.9
A:Map position: 2
A:Insertions: 160/3; 187/1; 243/1; 580/3; 607/3; 705/3; 789/1; 864/3; 940/3; 974/3; 1064/2;

Alignment Scores:
Pred. No.: 8.21 Length: 2329
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-19 (1-714) x T28125 (1-2329)

OY 412 AGATTTTAACTCAGTAATTTTA 435
|||||
DB 1632 ArgPheLeuIleSerLeuIleLeu 1639

RESULT 10
S34538
hypothetical protein 65 (rps18 3' region) - *Euglena gracilis* chloroplast
C:Species: chloroplast *Euglena gracilis*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Sep-1999
C:Accession: S34538; S34905
R:Halliack, R.B.; Hong, L.; Dräger, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielmann,
submitted to the EMBL Data Library, January 1993
A:Description: The complete sequence of the *Euglena gracilis* chloroplast genome (tentative)
A:Reference number: S34494
A:Accession: S34538
A:Molecule type: DNA
A:Residues: 1-65 <HAL1>
A:Cross-references: EMBL:X70810; NID:9415327; PIDN:CAA50118.1; PID:9415774
R:Halliack, R.B.; Hong, L.; Dräger, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spielmann,
Nucleic Acids Res. 21, 3537-3544, 1993
A:Title: Complete sequence of *Euglena gracilis* chloroplast DNA.
A:Reference number: S34862; MUID:93347989; PMID:8346031
A:Accession: S34905

A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-65 <HAL2>
A:Cross-references: EMBL:X70810; NID:9415327; PIDN:CAA50118.1; PID:9415774
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Genetics:
A:Genome: chloroplast
C:Superfamily: conserved hypothetical protein ycf9
C:Keywords: chloroplast

Alignment Scores:
Pred. No.: 140 Length: 65
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-19 (1-714) x S34538 (1-65)

OY 114 TTTTCTAGTTTTCATATA 94
|||||
DB 51 PhePheLeuValPheLeuIle 57

RESULT 11
S21595
19 heavy chain V region (7F2) - mouse
C:Species: *Mus musculus* (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S21595
R:Kaartinen, M.
submitted to the EMBL Data Library, May 1992
A:Reference number: S21591
A:Accession: S21595
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-86 <KAA>
A:Cross-references: EMBL:X66458; NID:951722; PIDN:CAA47073.1; PID:951723
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Alignment Scores:
Pred. No.: 136 Length: 86
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-19 (1-714) x S21595 (1-86)

OY 680 TATGTAATATGCTGGG 660
|||||
DB 76 TyrGlyAsnTyrGlyLeuGly 82

RESULT 12
E84075
hypothetical protein BH3405 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E84075
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E84075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; NID:910175792; PIDN:BAB07124.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3405

Alignment Scores:

Pred. No.:	133	Length:	105
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.21%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x E84075 (1-105)

OY 624 TTACGTCCTAATTTATTC 644
DB 11 LeuLeuPheLeuLeuPhe 17

RESULT 13

C70342
hypothetical protein ag_465 - Aquifex aeolicus

C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70342
R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70342
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-113 <AOP>
A:Cross-references: GB:AE000691; NID:g2983118; PIDN:AAC06729.1; PID:g2983125; GB:AE00065
A:Experimental source: strain VFS
C:Gene: ag_465
C:Superfamily: Aquifex aeolicus hypothetical protein ag_465

Alignment Scores:

Pred. No.:	132	Length:	113
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.23%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x C70342 (1-113)

OY 428 AGTGAGATTAATAATCTAAAA 408
DB 59 SerGluIleLysAsnLeuLys 65

RESULT 14

T30769
hypothetical protein BL - vaccinia virus (strain Ankara)

N:Alternate names: putative 13.7k protein
C:Species: vaccinia virus
A:Variety: strain Ankara
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R:Antoine, G.; Schellinger, F.; Falkner, F.G.; Dörner, F.
submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A:Reference number: Z20877
A:Accession: T30769
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-120 <ANT>
A:Cross-references: EMBL:U94848; PIDN:AAB96403.1
A:Experimental source: strain Ankara
C:Gene: MVA008L
A:Note: MVA008L

Alignment Scores:

Pred. No.:	131	Length:	120
------------	-----	---------	-----

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-19 (1-714) x T30769 (1-120)

OY 374 AAAATCTGTTCTGATAGCA 394
DB 2 LysIleLeuPheLeuIleAla 8

RESULT 15

F72150
B6L protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: F72150
R:Shchelkunov, S.N.; Tsvetnin, A.V.; Gutorov, V.V.; Safonov, P.F.; Massung, R.F.; Lo
submitted to Genbank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m
A:Reference number: A72150
A:Accession: F72150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <SHC>
A:Cross-references: GB:Y16780; NID:95830555; PIDN:CAB54600.1; PID:e1542556; PID:95830
A:Experimental source: strain Garcia-1966
C:Gene: B6L
A:Gene: B6L

Alignment Scores:

Pred. No.:	130	Length:	126
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.21%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-19 (1-714) x F72150 (1-126)

OY 374 AAAATCTGTTCTGATAGCA 394
DB 2 LysIleLeuPheLeuIleAla 8

Search completed: January 14, 2003, 17:40:35
Job time : 29.925 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 11.0253 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992a-19

Perfect score: 218
Sequence: 1 cgcagagaaagattatttca.....cacagtgttgccaattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+n2p-model -DEV-xip
-O/cgnt_1/USPTO_epool/US09835992/Runat_14012003_161525_4014/app-query.fasta.1.3932
-DB-SwissProt_40 -QFMT-fastan -SUFFIX-oligo.rsp -MIMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cd1
-LIST=45 -DOCMATCH=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFM-plc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR-US09835992.ccgnt.1.1.77.frunat.14012003.161525.4014 -NCFU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-WANT-TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6 -FCGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Seqs	Query Match	Length	DB ID	Description
1	8	3.7	208	1	PAFR_MACMU
2	8	3.7	342	1	PAFR_HUMAN
3	8	3.7	393	1	P306_MYGE
4	8	3.7	3178	1	YS89_CAREL
5	7	3.2	65	1	YCF9_EUGCR
6	7	3.2	113	1	Y465_AQUAE
7	7	3.2	164	1	2SS1_ARATH
8	7	3.2	164	1	2SS3_ARATH
9	7	3.2	180	1	YPKW_THRCU
10	7	3.2	194	1	YPMH_BACSU
11	7	3.2	201	1	YAHM_ECOLI
12	7	3.2	214	1	Y540_AQUAE
13	7	3.2	245	1	YCEI_RAT
14	7	3.2	267	1	Y494_PICST
15	7	3.2	268	1	Y494_PICST
16	7	3.2	278	1	HBD_DEIRA
17	7	3.2	278	1	Y110_MYGE
18	7	3.2	307	1	NADA_METJA

C	19	7	3.2	314	1	RLUC_BUCAI	P57430 buchiera ap
	20	7	3.2	314	1	Y315_MYCPN	Q50362 mycoplasma
	21	7	3.2	326	1	VS09_ROTET	Q03874 equine rota
	22	7	3.2	326	1	VS09_ROTET	P11855 human rotav
	23	7	3.2	326	1	VS09_ROTET	P12476 rhesus rota
	24	7	3.2	328	1	VMSA_HPBDO	P03145 duck hepari
	25	7	3.2	345	1	CLT2_PIG	Q95n03 sus scrofa
	26	7	3.2	365	1	VMSA_HPBDO	P30029 duck hepari
	27	7	3.2	366	1	VMSA_HPBDO	P17194 duck hepari
	28	7	3.2	366	1	VMSA_HPBDO	P17195 duck hepari
	29	7	3.2	379	1	PANE_YEAST	P37821 malus domes
	30	7	3.2	380	1	CAPM_STRAU	P39862 saccharomyc
	31	7	3.2	387	1	Y626_YEAST	P33063 saccharomyc
	32	7	3.2	397	1	PKA_AQUAE	O66519 aquifex aeo
	33	7	3.2	405	1	PKA_AQUAE	P57985 pasteurella
	34	7	3.2	423	1	MTB5_NEIGO	O59605 nelsaeria q
	35	7	3.2	423	1	MTB5_NEIGO	P50182 nelsaeria q
	36	7	3.2	430	1	DIN7_YEAST	Q12086 saccharomyc
	37	7	3.2	433	1	ENO_VIBCH	Q93p05 vibrio chol
	38	7	3.2	436	1	HMDH_ARCFU	O28538 archaeoglob
	39	7	3.2	460	1	Y454_HAEIN	P44104 haemophilus
	40	7	3.2	473	1	Y41C_MALDO	P37821 malus domes
	41	7	3.2	527	1	Y099_CAREL	P41849 caenorhabdi
	42	7	3.2	560	1	YUK5_YEAST	P42948 saccharomyc
	43	7	3.2	580	1	YKPI_XENLA	O13066 xenopus lae
	44	7	3.2	678	1	GYR_MYCLE	O59533 mycobacteri
	45	7	3.2	681	1	VGP_MABVY	P35253 marburg vir

ALIGNMENTS

RESULT 1
ID PAFR_MACMU STANDARD: PRT: 208 AA.
AC P35366: 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Platelet activating factor receptor (PAF-R) (Fragment).
GN PAFR.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_Taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Benhal R.H., Debussche M.S., Olson M.S.;
RL Submitted (xxx-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR. A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIANE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: L07333; AAA68893.1;
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1; 1.
CC PROSITE: PS00237; G-PROTEIN_RECPT_FL_1; 1.
CC PROSITE: PS50262; G-PROTEIN_RECPT_FL_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Chemotaxis.
CC DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
CC FT

```

FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT TRANSMEM 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT TRANSMEM 114 133 4 (POTENTIAL).
FT TRANSMEM 134 155 5 (POTENTIAL).
FT TRANSMEM 156 184 6 (POTENTIAL).
FT TRANSMEM 185 205 7 (POTENTIAL).
FT TRANSMEM 206 >208 8 (POTENTIAL).
FT CARBOHD 169 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFD 90 173 BY SIMILARITY.
FT NON_TER 208
SQ SEQUENCE 208 AA; 23736 MW; 7985928F70B3C6A1 CRC64;

Alignment Scores:
Pred. No.: 11 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x PAFR_MACMU (1-208)
QY 117 TCTTTTTCCTAGTTTCTAATA 94
Db 194 SerpPheLeuValPheLeuLe 201

RESULT 2
PAFR_HUMAN STANDARD; PRT; 342 AA.
ID PAFR_HUMAN STANDARD; PRT; 342 AA.
AC P25105;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet activating factor receptor (PAF-R).
OS P1AFR OR PAFR.
GN P1AFR.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92028922; PubMed=1656963;
RA Ye R.D., Prossnitz E.R., Zou A., Cochran C.G.;
RT "Characterization of a human cDNA that encodes a functional receptor
RT for platelet activating factor."
RL J. Biol. Chem. 266:20400-20405(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Leukocyte;
RX MEDLINE=92041873; PubMed=1657923;
RA Nakamura M., Honda Z., Izumi T., Sakanaka C., Mutoh H., Minami M.,
RT "Molecular cloning and expression of platelet-activating factor
RT receptor from human leukocytes."
RL J. Biol. Chem. 266:20400-20405(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250505; PubMed=1374385;
RA Kunz D., Gerard N.P., Gerard C.;
RT "The human leukocyte platelet-activating factor receptor. cDNA
RT cloning, cell surface expression, and construction of a novel
RT epitope-bearing analog."
RL J. Biol. Chem. 267:9101-9106(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347886; PubMed=1322356;
RA Seyfried C.E., Schweickart V.L., Godiska R., Gray P.W.;
RT "The human platelet-activating factor receptor gene (PTAFR) contains
RT no introns and maps to chromosome 1."
RL Genomics 13:832-834(1992).

```

```

RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93112021; PubMed=1281995;
RA Sugimoto T., Tsuchimochi H., McGregor C.G., Mutoh H., Shimizu T.,
RT Kurachi Y.;
RT "Molecular cloning and characterization of the platelet-activating
RT factor receptor gene expressed in the human heart."
RL Biochem. Biophys. Res. Commun. 189:617-624(1992).
RN [6]
RP SEQUENCE FROM N.A.
RA Behal R.H., Debysere M.S., Olson M.S.;
RL Submitted (XX-1992) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=93192035; PubMed=8383507;
RA Chase P.B., Halonen M., Regan J.W.;
RT "Cloning of a human platelet-activating factor receptor gene:
RT evidence for an intron in the 5'-untranslated region."
RL Am. J. Respir. Cell Mol. Biol. 8:240-244(1993).
CC -1- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR. A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTILE AND HYPOTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M80436; AAA60001.1; -
DR EMBL; M76674; AAA60002.1; -
DR EMBL; D10202; BAA01050.1; -
DR EMBL; M88177; AAA60214.1; -
DR EMBL; S52624; AAB24695.2; -
DR EMBL; I07334; AAA60108.1; -
DR EMBL; S56386; AAB25755.1; -
DR PIR; JH0479; JH0479.
DR PIR; A40191; A40191.
DR PIR; A41079; A41079.
DR Gene; HGNC:9582; PTAFR.
DR MIM; 173393;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_REC_P1_1;
DR PROSITE; PS00262; G_PROTEIN_REC_P2_1;
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Chemotaxis;
KW Polymorphism.
FT TRANSMEM 1 16
FT DOMAIN 17 38 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 39 54 1 (POTENTIAL).
FT TRANSMEM 55 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 91 2 (POTENTIAL).
FT TRANSMEM 92 113 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 114 133 3 (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT TRANSMEM 156 184 5 (POTENTIAL).
FT TRANSMEM 185 205 6 (POTENTIAL).
FT TRANSMEM 206 233 7 (POTENTIAL).
FT TRANSMEM 234 254 8 (POTENTIAL).
FT TRANSMEM 255 276 9 (POTENTIAL).
FT TRANSMEM 277 296 10 (POTENTIAL).
FT TRANSMEM 297 342 11 (POTENTIAL).
FT DISULFD 90 173 CYTOPLASMIC (POTENTIAL).
FT CARBOHD 169 173 BY SIMILARITY.
FT VARIANT 224 224 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT A -> D (IN DBSNP:5938).
FT /FTID=VAR_011851.

```


FT VARIANT 338 338 N -> S (IN DBSNP:5939).
 FT CONFLICT 28 28 /FTID=VAR_011852.
 FT CONFLICT 66 66 L -> P (IN REF. 6).
 FT CONFLICT 95 95 F -> L (IN REF. 6).
 FT CONFLICT 227 228 C -> R (IN REF. 6).
 FT CONFLICT 227 228 KR -> TG (IN REF. 4).
 FT CONFLICT 247 247 KR -> TT (IN REF. 6).
 FT CONFLICT 247 247 P -> A (IN REF. 6).
 FT CONFLICT 316 316 K -> N (IN REF. 5).
 SQ SEQUENCE 342 AA; 39203 MW; 890073C9EBA9228 CRC64;

Alignment Scores:

Prod. No.: 9.88 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x PAFR_HUMAN (1-342)

Qy 117 TCTTTTCTAGTTTCTATA 94

Db 194 SerpheeValpheeulle 201

RESULT 3

Y306_MYCGE STANDARD; PRT; 393 AA.
 ID Y306_MYCGE
 AC P47548;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG306.
 GN MG306.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 NC NCBI_TaxID=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischman R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.F., Dougherty B.A., Bort K.F., Hu P.C., Lucher T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U93711; AAC71528.1; -
 DR TIGR: MG306;
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 56 76 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 131 151 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 FT TRANSMEM 253 273 POTENTIAL.
 FT TRANSMEM 289 309 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 SQ SEQUENCE 393 AA; 45750 MW; D740BDA979EC364A CRC64;

Alignment Scores:

Prod. No.: 9.6 Length: 393
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservat: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x Y306_MYCGE (1-393)

Qy 117 TCTTTTCTAGTTTCTATA 94

Db 19 SerpheeValpheeulle 26

RESULT 4

YS89_CAEEL STANDARD; PRT; 3178 AA.
 ID YS89_CAEEL
 AC 009624; 009625; 0969D4;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK945.9 in chromosome II.
 GN ZK945.9/ZK945.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
 OC Rhabditidae; Pelodermidae; Caenorhabditis.
 NC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilkinson-Sproat J.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: Z48544; CAB70192.1; -
 DR EMBL: Z48582; CAB70192.1; JOINED.
 DR EMBL: Z48582; CAB70201.1; -
 DR EMBL: Z48544; CAB70201.1; JOINED.
 DR WormPep: ZK945.9; CE25697.
 DR InterPro: IPR002111; Cat_channel_TripL.
 DR InterPro: IPR001024; Lipoxigenase_LH2.
 DR InterPro: IPR000636; M-channel_nlg.
 DR InterPro: IPR000203; PKD_cys_rich.
 DR Pfam: PF00520; Ion_trans; 1.
 DR Pfam: PF01477; PLAT; 1.
 DR Pfam: PF01825; GPS; 1.
 DR SMART: SM00303; GPS; 1.
 DR SMART: SM00308; LH2; 1.
 KW Hypothetical protein; Transmembrane.
 FT DOMAIN 266 1196 SER/THR-RICH.
 FT DOMAIN 1105 1241 GLY/SER-RICH.
 FT DOMAIN 2071 2120 GPS.
 FT TRANSMEM 13 30 POTENTIAL.
 FT TRANSMEM 51 73 POTENTIAL.
 FT TRANSMEM 2139 2161 POTENTIAL.
 FT TRANSMEM 2348 2367 POTENTIAL.
 FT TRANSMEM 2390 2412 POTENTIAL.
 FT TRANSMEM 2451 2468 POTENTIAL.
 FT TRANSMEM 2483 2505 POTENTIAL.
 FT TRANSMEM 2567 2589 POTENTIAL.
 FT TRANSMEM 2836 2858 POTENTIAL.

FT TRANSMEM 2939 2961 POTENTIAL.
 FT TRANSMEM 2976 2998 POTENTIAL.
 FT TRANSMEM 3038 3060 POTENTIAL.
 SQ SEQUENCE 3178 AA: 344726 MW: 78239436D03666CD CRC64;

Alignment Scores:

Pred. No.: 6.17 Length: 3178
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.67% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x YS89_CABEL (1-3178)

OY 412 AGATTTTAACTGACTAAATTTTA 435

DB 2481 ArgPheLeuIleSerLeuIleLeu 2488

RESULT 5

YCF9_EUGGR STANDARD: PRT: 65 AA.

AC P32095;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 7.6 kDa protein ycf9 (ORF 65).

GN YCF9.

OS Euglena gracilis.

OC Chloroplast.

CC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.

OX NCBI_TaxID=3039;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z;

RX MEDLINE=93347989; PubMed=8346031;

RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Montfort A.,

Orat B., Spielmann A., Stutz E.,

"Complete sequence of Euglena gracilis chloroplast DNA.";

Nucleic Acids Res. 21:3537-3544(1993).

RL -1 SIMILARITY: BELONGS TO THE YCF9 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC -----

DR EMBL: Z11874; -; NOT_ANNOTATED_CDS.

DR EMBL: X70810; CAAS0118.1; -.

DR PIR: S34538; S34538.

DR InterPro: IPR002644; Ycf9_struct.

DR Pfam: PF01737; YCF9; 1.

DR Prodom: PD004770; Ycf9_struct; 1.

KM Chloroplast; Hypothetical protein.

SQ SEQUENCE 65 AA: 7586 MW: 67DE2359D7A0F771 CRC64;

Alignment Scores:

Pred. No.: 138 Length: 65
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.23% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x YCF9_EUGGR (1-65)

OY 114 TTTTCTAGTTTTCATATA 94

DB 51 PhePheLeuValPheLeuIle 57

RESULT 6

ID Y465_AQUAE STANDARD: PRT: 113 AA.

AC 066767;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Hypothetical protein AQ_465.

GN AQ_465.

OS Aquifex aeolicus.

CC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;

OC Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,

Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

"The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus.";

RL Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib-sib.ch).

CC -----

DR EMBL: AE000691; AAC06729.1; -.

DR Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 7 29 POTENTIAL.

SQ SEQUENCE 113 AA: 13252 MW: F402B9193B2A522 CRC64;

Alignment Scores:

Pred. No.: 123 Length: 113
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.23% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x Y465_AQUAE (1-113)

OY 428 AGGAGATTAATAAATCTAA 408

DB 59 SerGluIleLysAsnLeuLys 65

RESULT 7

ID 2S51_ARATH STANDARD: PRT: 164 AA.

AC P15457;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 2S seed storage protein 1 precursor (2S albumin storage protein)

GN (NMWU2-2S albumin 1).

OS Arabidopsis thaliana (Mouse-ear cress).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-73 AND 84-162.

RC STRAIN=cv. C24;

RA Krebbers E., Herdies L., de Clercq A., Seurlink J., Leemans J.,

van Damme J., Segura M., Gheysen G., van Montagu M.,

Vandekerckhove J.;

"Determination of the processing sites of an Arabidopsis 2S albumin

RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Conceicao A.D.S., Krebbers E.;
 RN Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wandut R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delany M., Puigdemont P., Watson M., Schmidtmann T.,
 RA Reichert B., Portelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohlebel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Biham L., Robben J.,
 RA Van der Schueren J., Glynoprez B., Chuang Y.-J., Vandenbusche F.,
 RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,
 RA Weitsengger T., Both G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirke W.,
 RA Moollman P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysaert C., Gleen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., Mcay K., Mayes R.,
 RA Pettit A., Rajandran M.A., Lynch M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Faltmann B., Granderath R., Dauner D., Hertzl A.,
 RA Neumann S., Argilou A., Vitale D., Lagouri R., Pirvandi E.,
 RA Messenc O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schenbl S., Hiller R., Schmidt W., Lecharry A., Aubourg S.,
 RA Cheffor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Holjen L., Schwarz S., Scholler P., Heber S., Frances P., Bietke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dehile N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheel P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latrille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramon J., Fulton L., Mardis E., Dante M., Papin K., Hillier L.,
 RA Nelson J., Splich J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Bergtholt A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matcero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Grant S., Shohdy N., Hasegawa A., Hamed A., Iodhi M., Johnson A.,
 RA Chen E., Marra M., Martensen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RL thaliana.";
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THIS IS THE MOST ABUNDANT ISOFORM OF 2S ALBUMIN IN
 CC ARABIDOPSIS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M2032; AAA32743.1; -.

DR EMBL; 224745; CAAB0870.1; -;
 DR EMBL; A13820; CAA01132.1; -;
 DR EMBL; AL035680; CAB38844.1; -;
 DR EMBL; A161566; CAB79569.1; -;
 DR PIR; JAO161; NMWU1;
 DR PIR; S34676; S34676;
 DR InterPro: IPR003612; AAI;
 DR InterPro: IPR000617; Napin.
 DR InterPro: IPR001768; Try/amyL_inhbr.
 DR Pfam: PF00234; tryr_alpha_amyL; 1.
 DR PRINTS; PR00496; NAPIN.
 DR ProDom; PD002498; Napin; 1.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Albumin; signal; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 37
 FT CHAIN 38 73
 FT PROPEP 74 83
 FT CHAIN 84 162
 FT PROPEP 163 164
 SQ SEQUENCE 164 AA; 19014 MW; 2BF28CB47AD9832B CRC64;
 Alignment Scores:
 Pred. No.: 114
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.21%
 DB: 1
 Gaps: 0
 Length: 164
 Matches: 7
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0
 US-09-835-992a-19 (1-714) x 25S1_ARATH (1-164)
 QY 496 TTGGCATTTGCTTCTACTT 516
 Db 12 Leubaleucysphreuleu 18
 RESULT 8
 25S1_ARATH STANDARD; PRT; 164 AA.
 ID 25S1_ARATH
 AC P15459;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 2S seed storage protein 3 precursor (2S albumin storage protein)
 DE (NMWU2-2S albumin 3).
 GN AT253 OR AT4G27160 OR T24A18.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Krebbers E., Herdies L., de Clercq A., Seurinck J., Leemans J.,
 RA van Damme J., Segura M., Gheysen G., van Montagu M.,
 RA Vandekerckhove J.;
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Conceicao A.D.S., Krebbers E.;
 RN Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Columbia;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wandut R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,

RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichelt B., Portetelle D., Perez-Alonso M., Bottry M., Bancroft I.,
RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Roben J.,
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
RA Braken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weltenegeer T., Bothe G., Ransperger U., Hilbert H., Braun M.,
RA Holter E., Brandt A., Peters S., van Staveren M., Dirse W.,
RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koelter P.,
RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIay K., Mayes R.,
RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandeth K., Dauner D., Herzl A.,
RA Neumann S., Argitrou A., Vitale D., Liguori R., Piravandi E.,
RA Masenet O., Outgley F., Schmidt W., Lechamy A., Aubourg S.,
RA Schnabl S., Hiller R., Schmidt W., Lechamy A., Aubourg S.,
RA Chedot F., Cooke R., Berger C., Montfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekton M., Murray J., Sheet P., Cordes M., Abu-Threidh J.,
RA Storeking T., Kallick J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mux P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speith J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Bergthoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Mateo A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.,
RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RA thaliana.";
RT Nature 402:769-777(1999).
RL [4]
RN
RP SEQUENCE OF 103-164 FROM N.A.
RC STRAIN-CV. Columbia;
RA Raynal M., Gréillet F., Laugie M., Meyer Y., Cooke R., Delseny M.;
RL Submitted (OCT1992) to the EMBL/Genbank/DBJ databases.
CC
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC
CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY DISULFIDE BONDS.
CC
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M22035; AAA32745.1;
DR EMBL: Z24744; CAB80868.1;
DR EMBL: AL035680; CAB38846.1;
DR EMBL: AL161566; CAB79571.1;
DR EMBL: Z17580; CAB79001.1;
DR PIR: JAO163; NMM03.
DR PIR: S34674; S34674.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000617; Napin.
DR InterPro: IPR001768; try/amy1_inhbr.
DR Pfam: PF00234; try_alpha_amy1; 1.
DR PRINTS: PR00496; NAPIN.
DR ProDom: PD002498; Napin; 1.

DR SMART: SM00499; AAI; 1.
KW Seed storage protein; Albumin; Signal; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 37
FT CHAIN 38 72 2S SEED STORAGE PROTEIN 3 SMALL SUBUNIT
FT (BY SIMILARITY).
FT PROPEP 73 81
FT CHAIN 82 164
FT 2S SEED STORAGE PROTEIN 3 LARGE SUBUNIT
SQ SEQUENCE 164 AA; 18762 MW; C9BEB6718549F248 CRC64;
(BY SIMILARITY).
Alignment Scores:
Pred. No.: 114 Length: 164
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-19 (1-714) x 2SS3_ARATH (1-164)
OY 496 TTGCGATGCTGTTCTTACTT 516
|||||
DB 12 LeuAlaLeuCySPheLeuLeu 18
RESULT 9
ID YPKW_THECU STANDARD; PRT; 180 AA.
AC P496594;
DT 01-FEB-1996 (rel. 33, Created)
DT 01-FEB-1996 (rel. 33, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Hypothetical protein in pKWA 5' region (ORF1) (Fragment).
OS Thermomonospora curvata.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptosporanginae; Thermomonosporaceae;
OC Thermomonospora;
OX NCBI_TaxID=2020;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CM 3352;
RX MEDLINE=96200125; PubMed=8631732;
RA Tanda L., Tichy P., Spizek J., Petricek M.;
RT serine/threonine protein kinase and WD-repeat domains.";
RL J. Bacteriol. 178:1487-1489(1996).
CC
CC -1- SIMILARITY: BELONGS TO THE CDAR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF115313; NOT_ANNOTATED_CDS.
DR InterPro: IPR000712; Bcl2_BH.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 180 AA; 19880 MW; 7399D55F9E204D40 CRC64;
Alignment Scores:
Pred. No.: 112 Length: 180
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.23% Indels: 0
DB: 1 Gaps: 0
US-09-835-992a-19 (1-714) x YPKW_THECU (1-180)
OY 76 AATTCAGTGAGATATAGGATA 56

```

Db      132 AsnSerValArgTyrGlyIle 138
RESULT 10
YPBH_BACSU          STANDARD;          PRT;          194 AA.
AC      PS0736;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Hypothetical protein ypbh.
GN      ypbh.
OS      Bacillus subtilis.
OC      Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX      NCBI_TaxID=1423;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=96349105; PubMed=8760912;
RA      Sorokin A.V., Azavedo V., Zumbstein E., Galleron N., Ehrlich S.D.,
RA      Serror P.;
RT      "Sequence analysis of the Bacillus subtilis chromosome region between
RT      the aera and kds loci cloned in a yeast artificial chromosome.";
RL      Microbiology 142:2005-2016(1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=168;
RX      MEDLINE=98044033; PubMed=9384377;
RA      Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA      Azevedo V., Berrero M.C., Bessieres P., Boletini A., Borchert S.,
RA      Bottaris R., Boutsier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA      Broutlier S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA      Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA      Denicourt F., Devine K.M., Dusterhoft A., Ehlich S.D., Emerson P.T.,
RA      Ertlan K.D., Errington J., Fabret C., Ferrari E., Fougere D.,
RA      Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA      Gilm S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA      Halberst G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA      Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA      Joris B., Karamata D., Kasehara Y., Kleier-Blanchard M., Klein C.,
RA      Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA      Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA      Lee S.M., Levine A., Liu H., Masuda S., Meisel C., Medigue C.,
RA      Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Nodack M.,
RA      Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA      Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
RA      Presseau E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA      Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA      Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA      Sekiguchi J., Skowaska A., Soror S.J., Serror P., Shin B.S., Soldo B.,
RA      Sokolich A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA      Takeuchi V., Yamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA      Tosaio V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,
RA      Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA      Winters P., Wipit A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA      Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT      "The complete genome sequence of the Gram-positive bacterium Bacillus
RT      subtilis.";
RL      Nature 390:249-256(1997).
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: LA7648; AAC83952.1; -
CC      EMBL: 299115; CAB14213.1; -
CC      EMBL: 299116; CAB14229.1; -
CC      Subtilisin; BGI1634; ypbh.
KW      Hypothetical protein; Complete proteome.

```

```

SQ      SEQUENCE      194 AA;      22157 MW;      B27C6ABE4EA0745C CRC64;
Alignment Scores:
Pred. No.:      110
Score:      7.00
Percent Similarity:      100.00%
Best Local Similarity:      100.00%
Query Match:      3.23%
DB:      1
Gaps:      0
US-09-835-992a-19 (1-714) x YPBH_BACSU (1-194)
OY      317 AAATTAATAATATTTTACCA 297
Db      10 Lysileysilephenethr 16
RESULT 11
YAHB_ECOLI          STANDARD;          PRT;          201 AA.
ID      YAHB_ECOLI
AC      P77736;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Hypothetical ANK-repeats protein yahb.
GN      YAHB OR B0318.
OS      Escherichia coli.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=K12 / MG1655;
RX      MEDLINE=97426617; PubMed=9278503;
RA      Blattner F.R., Plunkett G. Iii, Bloch C.A., Perna N.T., Burland V.,
RA      Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA      Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA      Mau B., Shao Y.;
RT      "The complete genome sequence of Escherichia coli K-12.";
RL      Science 277:1453-1474(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA      Federpiet N., Hymen R., Kalman S., Komp C., Kurdi O., Lew H., Lin D.,
RA      Natch A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC      1- SMITABILITY: CONTAINS 6 ANK REPEATS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AE000139; AAC73421.1; -
CC      EMBL: U73857; AAB18044.1; -
CC      Ecogene: EG13588; yahb.
CC      InterPro: IPR002110; ANK.
CC      Pfam: PF00023; ANK; 6.
CC      SMART: SM00248; ANK; 2.
CC      PROSITE: PS50088; ANK_REPEAT; 3.
CC      PROSITE: PS50297; ANK_REPEAT_REGION; 1.
CC      KW      Hypothetical protein; Repeat; ANK repeat; Complete proteome.
CC      -----
CC      REPEAT      5      34      ANK 1.
CC      REPEAT      6      37      ANK 2.
CC      REPEAT      7      100      ANK 3.
CC      REPEAT      8      134      ANK 4.
CC      REPEAT      9      172      ANK 5.
CC      REPEAT      10      201      ANK 6.
CC      SEQUENCE      201 AA;      21687 MW;      A90244CEB892EBBD CRC64;
Alignment Scores:

```

```

Pred. No.: 109 Length: 201
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x YAHU_ECOLI (1-201)
OY 326 GCCTGTGTCACGCTCGATC 346
|||||
DB 53 AlacysValGlnAlaLeuile 59

RESULT 12
Y540_AOUAE STANDARD; PRT; 214 AA.
AC 066819;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AO_540.
GN AO_540.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificaceae; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJAYI M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -1 SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1 SIMILARITY: BELONGS TO THE UPR0056 (MARC) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: AE000694; AAC06780.1;
CC DR InterPro: IPR002771; MARC.
CC DR Pfam: PF01914; MARC.
CC DR TIGRPFAM: TIGR00427; UPR0056; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SQ SEQUENCE 214 AA; 23572 MW; DDIA9947A5DE931D CRC64;

Alignment Scores:
Pred. No.: 108 Length: 214
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.21% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x Y540_AOUAE (1-214)
OY 618 GCGATTTTACTGTTCTATT 638
|||||
DB 86 GlyIleLeuLeuPheLeuile 92

```

```

RESULT 13
FCEL_RAT STANDARD; PRT; 245 AA.
ID FCEL_RAT
AC P12371;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin epsilon receptor alpha-subunit precursor
DE (FCER1) (IGE FC receptor, alpha-subunit) (FC-epsilon RI-alpha).
GN FCER1A OR FCEL1A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88024987; PubMed=2959318;
RA Kinet J.-P., Metzger H., Hakim J., Kochan J.;
RT "A cDNA presumably coding for the alpha subunit of the receptor
RT with high affinity for immunoglobulin E."
RL Biochemistry 26:4605-4610(1987).
RN [2]
RP REVISIONS.
RA Kochan J.;
RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Test cells;
RX MEDLINE=88158102; PubMed=2964640;
RA Shimizu A., Tepler I., Benley P.N., Berenstein E.H., Siraganian R.P.,
RA Leder P.;
RT "Human and rat mast cell high-affinity immunoglobulin E receptors:
RT characterization of putative alpha-chain gene products."
RL Proc. Natl. Acad. Sci. U.S.A. 85:1907-1911(1988).
RN [4]
RP SEQUENCE OF 21-245 FROM N.A.
RX MEDLINE=88289772; PubMed=2969594;
RA Liu F.-T., Albrandt K., Robertson M.W.;
RT "cDNA heterogeneity suggests structural variants related to the high-
RT affinity Ige receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 85:5639-5643(1988).
CC -1 FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1 SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: M17153; AAA42045.1;
CC DR EMBL: J03606; AAA41582.1;
CC DR EMBL: M21622; AAA41146.1;
CC DR PIR: A27116; A27116.
CC DR PIR: A30154; A30154.
CC DR PIR: C31327; C31327.
CC DR HSSP: P12319; IALS.
CC DR InterPro: IPR003599; I9.
CC DR InterPro: IPR003006; I9_MHC.
CC DR Pfam: PF00047; I9; 2.
CC DR SMART: SM00409; I9; 2.
CC KW Ige-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;
KW

```

KM Immunoglobulin domain: Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 25
 FT TRANSMEM 205 204
 FT TRANSMEM 205 223
 FT DOMAIN 224 225
 FT DOMAIN 42 98
 FT DOMAIN 123 181
 FT DISULFID 130 174
 FT CARBOHYD 52 52
 FT CARBOHYD 53 53
 FT CARBOHYD 58 58
 FT CARBOHYD 65 65
 FT CARBOHYD 123 123
 FT CARBOHYD 158 158
 FT CARBOHYD 167 167
 SQ SEQUENCE 245 AA: 27793 MW: A0B67DD363B72197 CRC64:

Alignment Scores:
 Pred. No.: 105 Length: 245
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.21% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x PCE1_RAT (1-245)

OY 83 AATAACATTAGTATTGAAAA 103
 DB 157 AsnAsnIleSerIleArgLys 163

RESULT 14

PYRF_PICST STANDARD: PRT: 267 AA.

AC P49344;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
 DE (OMPDCase) (OMPDCase) (Uridine 5'-monophosphate synthase) (UMP synthase).
 GN URA3.
 OS Pichia stipitis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4924;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 58785 / CBS 6054;
 RX MEDLINE=95110115; PubMed=7811063;
 RA Yang V.W., Marks J.A., Davis B.P., Jeffries T.W.;
 RT High efficiency transformation of Pichia stipitis based on its URA3
 RT gene and a homologous autonomous replication sequence, ARS2.*;
 RL Appl. Environ. Microbiol. 60:4245-4254(1994).
 CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate -> UMP + CO(2).
 CC -1- PATHWAY: Pyrimidine biosynthesis: sixth (last) step.
 CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U08629; AAA65978.1; -.
 DR HSSP: P03962; IDOW.
 DR InterPro: IPR001754; OMPdecase.
 DR Pfam: PF00215; OMPdecase.1.
 DR Pfam: PF00215; OMPdecase.1.

DR PROSITE: PS00156; OMPDECASE.1.
 KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
 FT ACT_SITE 94 94 BY SIMILARITY.
 SQ SEQUENCE 267 AA: 29529 MW: D98361C398F49E06 CRC64:

Alignment Scores:
 Pred. No.: 103 Length: 267
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.23% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x PYRF_PICST (1-267)

OY 521 GAACCTAGCTAGACAAACAAT 501
 DB 80 GluIleSerLysLysHisAsn 86

RESULT 15

PYRF_CANTR STANDARD: PRT: 268 AA.

AC Q42771;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
 DE (OMPDCase) (OMPDCase) (Uridine 5'-monophosphate synthase) (UMP synthase).
 GN URA3.
 OS Candida tropicalis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M4;
 RX MEDLINE=98403409; PubMed=9734982;
 RA Su J.-H., Hsia J.-H., Chang M.-C.;
 RT Cloning and sequence analysis of the Candida tropicalis URA3 gene
 RT encoding orotidine-5'-phosphate decarboxylase.*;
 RL Curr. Microbiol. 37:210-213(1998).
 CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate -> UMP + CO(2).
 CC -1- PATHWAY: Pyrimidine biosynthesis: sixth (last) step.
 CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF040702; AAB96773.1; -.
 DR HSSP: P03962; IDOW.
 DR InterPro: IPR001754; OMPdecase.
 DR Pfam: PF00215; OMPdecase.1.
 KW PROSITE: PS00156; OMPDECASE.1.
 KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
 FT ACT_SITE 94 94 BY SIMILARITY.
 SQ SEQUENCE 268 AA: 29672 MW: A43C1D1CB24237 CRC64:

Alignment Scores:

Pred. No.: 103 Length: 268
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.23% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992a-19 (1-714) x PYRF_CANTR (1-268)

OY 521 GAACTAAGTAAGAAACACAAAT 501
|||||
Db 80 GluLeuSerLysLysHisAsn 86

Search completed: January 14, 2003, 17:29:07
Job time : 17.0253 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 50.2755 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992a-19
Perfect score: 218
Sequence: 1 cgcgcgaaaaagtcatttca.....cacagttgttcacattca 714

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE-frame+ n2p.model -DEV-xlp
-O/cgcn2.1/USPTO.spool/US09835992/runal.14012003.161526.4027/app.query.fasta.1.3932
-DB-SPTREMBL.21 -QWTF-fasten -SUFFIX-oligo.rspt -MINMATCH-0.1 -LOOPEL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human4.coi
-LIST-45 -DOCALLIGN-200 -THR_SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pio -NOR-ext -HEAPSIZ-500 -MINLEN-0 -MAXLEN-200000000
-USER-US09835992.ecgn.1.1.346.@runat.14012003.161526.4027 -NCPU-6 -ICPU-3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV-TIMEOUT-120
-MARN_TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YCAPOP-60 -YGAPEXT-60 -DELop-6 -DELEXT-7

Database :

1: SPTREMBL.21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_birchep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	3.7	123	17	0973p2
					0973p2 sulfolobus

C	2	8	3.7	123	17	0962w9	0962w9 sulfolobus
C	3	8	3.7	151	6	097662	097662 ovis aries
C	4	8	3.7	195	17	053004	053004 pyrococcus
C	5	8	3.7	280	5	09N5G4	09N5G4 caenorhabditis
C	6	8	3.7	342	6	09GK76	09GK76 capra hircus
C	7	8	3.7	342	6	09TMY5	09TMY5 bos taurus
C	8	8	3.7	342	5	017820	017820 caenorhabditis
C	9	8	3.7	352	17	08TWT5	08TWT5 methanobacterium
C	10	8	3.7	375	16	08XN39	08XN39 clostridium
C	11	8	3.7	376	16	092RV3	092RV3 rhizobium
C	12	8	3.7	413	17	0972W9	0972W9 sulfolobus
C	13	8	3.7	485	11	09JMT5	09JMT5 ratulus norv
C	14	8	3.7	485	11	09JPK1	09JPK1 mus musculus
C	15	8	3.7	512	16	09HVA7	09HVA7 pseudomonas
C	16	8	3.7	791	17	09YB89	09YB89 aeropyrum p
C	17	8	3.7	1115	4	09HDA3	09HDA3 homo sapien
C	18	8	3.7	1118	4	015426	015426 homo sapien
C	19	8	3.7	1183	16	0980H9	0980H9 mycoplasma
C	20	8	3.7	3178	5	0969D4	0969D4 caenorhabditis
C	21	7	3.2	55	10	041910	041910 arabidopsis
C	22	7	3.2	55	12	091FK1	091FK1 chilo tride
C	23	7	3.2	72	6	09GMR4	09GMR4 macaca fasc
C	24	7	3.2	85	5	09VDM5	09VDM5 drosophila
C	25	7	3.2	105	16	09K7F6	09K7F6 bacillus ha
C	26	7	3.2	106	2	047966	047966 herpetosiph
C	27	7	3.2	117	8	009313	009313 nesticus st
C	28	7	3.2	117	8	009314	009314 nesticus st
C	29	7	3.2	117	8	009317	009317 nesticus ba
C	30	7	3.2	117	8	009333	009333 nesticus sl
C	31	7	3.2	120	12	057169	057169 vaccinia vi
C	32	7	3.2	121	2	09AL17	09AL17 enterococcu
C	33	7	3.2	124	12	P87608	P87608 compox vitu
C	34	7	3.2	126	12	089067	089067 variola vir
C	35	7	3.2	139	8	095127	095127 neoneurus m
C	36	7	3.2	167	12	09WFA4	09WFA4 snow goose
C	37	7	3.2	167	12	09WFA7	09WFA7 snow goose
C	38	7	3.2	167	12	09WFB0	09WFB0 snow goose
C	39	7	3.2	167	12	09WFB4	09WFB4 snow goose
C	40	7	3.2	167	12	09WFB7	09WFB7 snow goose
C	41	7	3.2	167	12	090742	090742 duck hepati
C	42	7	3.2	167	12	089748	089748 duck hepati
C	43	7	3.2	167	12	089689	089689 duck hepati
C	44	7	3.2	181	8	0954U1	0954U1 panoploctri
C	45	7	3.2	183	10	08RWP7	08RWP7 arabidopsis

ALIGNMENTS

RESULT 1
0973p2 PRELIMINARY; PRT; 123 AA.
ID 0973p2
AC 0973p2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Putative cryptogene protein G4.
GN S70856.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankei A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.,
RT Complete genome sequence of an aerobic thermocacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.

RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000984; BAB65869.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 123 AA; 14153 MW; 4C04001E12BD02 CRC64;

Alignment Scores:

Pred. No.:	35.6	Length:	123
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992a-19 (1-714) x Q973P2 (1-123)

Oy 120 TTATCTTTTCTAGTTTCTA 97
 |||||
 Db 106 LeuserPheleuValPheleu 113

RESULT 2

O96ZM9 PRELIMINARY; PRT; 123 AA.

AC O96ZM9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN ST1718.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JCM 10545 / 7;
 RX PubMed-11572479;
 RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 Sakine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
 Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Ogunchi A.,
 Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermocacidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL; AP000987; BAB66804.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 123 AA; 14210 MW; DAAD4AC9B8A5826F CRC64;

Alignment Scores:

Pred. No.:	35.6	Length:	123
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992a-19 (1-714) x Q96ZM9 (1-123)

Oy 120 TTATCTTTTCTAGTTTCTA 97
 |||||
 Db 106 LeuserPheleuValPheleu 113

RESULT 3

O97662 PRELIMINARY; PRT; 151 AA.

AC O97662;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Platelet activating factor receptor (Fragment).
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=LUNG;
 RA Sander F.C., Ibe O.B., Raj J.U.;
 RT "Developmentally Regulated Expression of Platelet Activating Factor
 RT Receptor in Ovine Lung."
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF099674; AAC77459.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.

DR PRINTS: PR00237; GPCRHHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECPE_F1_2; 1.

KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 151 151
 SQ SEQUENCE 151 AA; 17001 MW; 31EC4F2006AF951C CRC64;

Alignment Scores:

Pred. No.:	34.1	Length:	151
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992a-19 (1-714) x O97662 (1-151)

Oy 117 TCTTTTCTAGTTTCTAATA 94
 |||||
 Db 138 SerPheleuValPheleu 145

RESULT 4

O59004 PRELIMINARY; PRT; 195 AA.

AC O59004;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein PH1301.
 GN PH1301.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed-9679194;
 RA Kawarabayasi Y., Sawada M., Horikawa H., Hakiwa Y., Hino Y.,
 Yamamoto S., Sakine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohluku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Ogunchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masui Y., Shizuya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 DR EMBL; AP000006; BAA30405.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 195 AA; 22883 MW; C1075554545837CC CRC64;

Alignment Scores:

Pred. No.:	32.2	Length:	195
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	17	Gaps:	0

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8.00	28.5	342	8	0	0	0	0
100.00%							
100.00%							
3.69%							

US-09-835-992a-19 (1-714) x Q9T7Y5 (1-342)

OY 117 TCTTTTTCCTAGTTTCTAATA 94
 DB 194 SerpPhelValpHeuLeulle 201

RESULT 8

Q17820 PRELIMINARY: PRT: 343 AA.

AC 017820: 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE C08B6.9 protein.

GN C08B6.9.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Wilkinson J.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for investigating biology."

RL Science 282:2012-2018(1998).

DR EMBL: 272502; GAA96591.1; -.

DR InterPro: IPR000594; Th1F_domain.

DR Pfam: PF00899; Th1F; 1.

SO SEQUENCE 343 AA; 38738 MW; 2897C0F918E881B6 CRC64;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8.00	28.4	343	8	0	0	0	0
100.00%							
100.00%							
3.69%							

US-09-835-992a-19 (1-714) x Q17820 (1-343)

OY 646 CAGATTAATAATAGGACGATAA 623
 DB 23 GlnAsnLysIleArgAsnSerLys 30

RESULT 9

Q8T7Z5 PRELIMINARY: PRT: 352 AA.

AC 08T7Z5: 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Sulfate/molybdate ABC transporter, solute-binding protein.

GN MA0280.

OS Methanosaarcina acetivorans.

OC Archaea; Euryarchaeota; Methanococci; Methanosaeriales;

OC Methanosaeriales; Methanosaerinae; Methanosaerina.

OX NCBI_TaxID=2214;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=C2A / ATCC 35395 / DSM 2834;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8.00	28.3	352	8	0	0	0	0
100.00%							
100.00%							
3.69%							

US-09-835-992a-19 (1-714) x Q8XN39 (1-352)

OY 114 TTTTCTCCTAGTTTCTAATACTA 91
 DB 10 PhePhelValpHeuLeulleu 17

RESULT 10

Q8XN39 PRELIMINARY: PRT: 375 AA.

AC 08XN39: 01-MAR-2002 (TREMBLrel. 20, Created)

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)

DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Probable hexosyltransferase.

GN CPE0499.

OS Clostridium perfringens.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;

OC Clostridiales; Clostridiaceae; Clostridium.

OX NCBI_TaxID=1502;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=13 / TYPE A;

RC PubMed=11792842;

RA Shmizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,

RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;

RT "Complete genome sequence of Clostridium perfringens, an anaerobic flesh eater."

RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).

DR EMBL: AP003187; BAB80205.1; -.

DR InterPro: IPR001296; Glycosyltransf_1.

DR Pfam: PF00534; Glycosyltransf_1; 1.

KW Transferrase; Complete proteome.

SO SEQUENCE 375 AA; 43028 MW; AB7626E3FF9E7B56 CRC64;

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8.00	27.9	375	8	0	0	0	0
100.00%							
100.00%							
3.69%							

US-09-835-992a-19 (1-714) x Q8XN39 (1-375)

OY 428 AGTGAGTTAAATCTAATAATT 405
 DB 10 PhePhelValpHeuLeulleu 17

Db 47 Serulilelysaenleu54

RESULT 11

ID 092RV3 PRELIMINARY; PRT: 376 AA.

AC 092RV3:

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Putative permease protein.

GN R00738 OR SMC00793.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021.

RA MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Boleard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,

RA Pohl T., Portetelle P., Puehler A., Purnelle B., Ransperger U.,

RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."

RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

DR EMBL: AL591784; CAC45310.1; "

DR InterPro: IPR000560; HIsAc-phsptase.

DR InterPro: IPR002548; UPF0118.

DR Pfam: PF01594; UPF0118; 1.

DR PROSITE: PS00778; HIs-ACID-Phosphat_2; UNKNOWN_1.

KW Complete proteome.

SQ SEQUENCE 376 AA; 40673 MW; F119E2DFBFC87 CRC64;

Alignment Scores:

Pred. No.:	27.9	Length:	376
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992a-19 (1-714) x 092RV3 (1-376)

OY 493 CTGTTGCATTGCTTCTACTT 516

Db 97 LeuLalaleuCyphenleu 104

RESULT 12

ID 0972W9 PRELIMINARY; PRT: 413 AA.

AC 0972W9:

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Putative D-hydantoinase.

GN S71020.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=111955;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / 7;

RX PubMed=11572479;

RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,

RA Sokino M., Baba S.-I., Ankel A., Kosugi H., Hosoyama A., Fukui S.,

RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,

RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,

RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,

RA Ohshima T., Kikuchi H.;

KW Complete genome sequence of an aerobic thermophilic

RT Crenarchaeon, Sulfolobus tokodaii strain7.;

RL DNA Res. 8:123-140(2001).

DR EMBL: AP000984; BAB66044.1; "

DR InterPro: IPR002195; Dihydroorotase.

DR Pfam: PF00744; Dihydroorotase; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 413 AA; 46760 MW; CE5B584E39710A9F CRC64;

Alignment Scores:

Pred. No.:	27.3	Length:	413
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992a-19 (1-714) x 0972W9 (1-413)

OY 517 AGTTCTCCCAAGGAAACTCTTA 540

Db 405 SerseProlysglyLeu 412

RESULT 13

ID 09JM15 PRELIMINARY; PRT: 485 AA.

AC 09JM15:

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Neuronal glutamine transporter.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RA MEDLINE=20127882; PubMed=10660562;

RA Varoqui H., Zhu H., Yao D., Ming H., Erickson J.D.;

RT "Cloning and Functional Identification of a Neuronal Glutamine

RT Transporter."

RL J. Biol. Chem. 275:4049-4054(2000).

DR EMBL: AF075704; AAF34240.1; "

DR InterPro: IPR002422; AA/rel-primase2.

DR Pfam: PF01490; AA_trans; 1.

SQ SEQUENCE 485 AA; 53846 MW; OCCDAD864513FC186 CRC64;

Alignment Scores:

Pred. No.:	26.4	Length:	485
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992a-19 (1-714) x 09JM15 (1-485)

OY 618 GGGATTTCAGTCCTAATTGA 641

Db 100 GlylleuLeuPheulleu 107

RESULT 14

ID 099PR1 PRELIMINARY; PRT: 485 AA.

AC 099PR1:

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Amino acid transporter Nat-2.

GN AA408026 OR NAT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=21316561; PubMed=11325958;
 RA Gu S., Roderick H.L., Camacho P., Jiang J.X.;
 RT "Characterization of an N-system Amino Acid Transporter Expressed in
 RL Retina and Its Involvement in Glutamine Transport.";
 DR J. Biol. Chem. 276:24137-24144(2001).
 DR EMBL: AF184240; AAG43433.2; -;
 DR MGD: MGI:2145895; AA408026.
 DR InterPro: IPR002422; AA/rel_pmease2.
 DR Pfam: PF01490; Aa_trans: 1;
 SQ SEQUENCE 485 AA; 53739 MW; 11636A57030F976B CRC64;

Alignment Scores:

Pred. No.:	26.4	Length:	485
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.67%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992A-19 (1-714) x Q9HVA7 (1-485)

QY 618 GCGATTTCCTGCTCTAATTGA 641

DB 100 GlyileLeuLeuPheLeuLeu 107

RESULT 15

Q9HVA7

ID Q9HVA7; PRELIMINARY; PRT; 512 AA.

AC Q9HVA7; 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)

DE Iron (III)-transport system permease HtrB.

GN HtrB OR PA4688.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Coltery L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen";

RL Nature 406:959-964(2000).

CC -!- FUNCTION: PROBABLY PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT

CC SYSTEM. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE

CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -!- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-

CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.

CC EMBL: AE004882; AAG08075.1; -;

DR InterPro: IPR000515; BPD_transp.

DR Pfam: PF00528; BPD_transp; 2.

DR PROSITE: PS00402; BPD_TRANS_P. INN_MEMBER: 1.

KW Transmembrane; Transport; Complete proteome.

SQ SEQUENCE 512 AA; 56176 MW; 9A21A90B3CDBF0A CRC64;

Alignment Scores:

Pred. No.:	26	Length:	512
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match:	3.67%	Indels:	0
DB:	16	Gaps:	0

US-09-835-992A-19 (1-714) x Q9HVA7 (1-512)

QY 493 CTCTGGCATGCTGCTTCTACTT 516

DB 230 LeuLeuAlaLeuLeuLeuLeu 237

Search completed: January 14, 2003, 17:36:49
 Job time : 54.2755 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 33.8408 Seconds

(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 212
Sequence: 1 atttaatttcataaac.....cttgatgacacagtgtc 687

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
Ygapop	60.0 , Ygapext 60.0
Fgapop	6.0 , Fgapext 7.0
Delop	6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+ -n2p -model -DEV-xip
-O-cg92.1/USPTO.epool/US09835992.r/unat.14012003.161525.3979/app_query.fasta.1.3932
-DB-A-Geneseq.101002 -ORFW-fastan -SUFFIX-olig.rag -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITs-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.ccl
-LIST=45 -DOCALLIN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFW=pio -MOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPTO-US09835992.ecgn.1.1.121.g/unat.14012003.161525.3979 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCAPOP=6 -FCAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database: A_Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	3.8	29	AAW40006	Peptide effecting
2	8	3.8	104	AAO11927	Human polypeptide
3	8	3.8	159	ABB31093	Peptide #3744 enco
4	8	3.8	159	ABB36288	Peptide #3794 enco
5	8	3.8	159	ABB21653	Protein #3652 enco
6	8	3.8	159	AAW57060	Human brain expres
7	8	3.8	159	AAW69450	Human bone marrow
8	8	3.8	159	AAW17284	Peptide #3718 enco
9	8	3.8	159	AAW27974	Peptide #3621 enco
10	8	3.8	159	AAW04971	Peptide #3653 enco
11	8	3.8	275	ABG39070	Human peptide enco
12	8	3.8	275	ABB09061	Thermus caldophilu
13	8	3.8	342	AAW25855	Human platelet act
14	8	3.8	342	AAW49550	Human platelet act
15	8	3.8	342	ABW56381	Non-endogenous hum
16	8	3.8	342	ABW77904	Amino acid sequenc
17	8	3.8	342	ABW77904	Amino acid sequenc
18	8	3.8	542	AAW70385	Yeast host cell pr
19	8	3.8	631	AAW56847	Arabidopsis thalia
20	8	3.8	631	AAW56846	Arabidopsis thalia
21	8	3.8	743	AAW56845	Caenorhabditis ele
22	8	3.8	2870	AAW55559	Caenorhabditis ele
23	8	3.8	3178	AAW55556	Caenorhabditis ele
24	7	3.3	15	ABW62591	Eubacterial DNA po
25	7	3.3	18	AAW92053	Modified type of e
26	7	3.3	18	AAW05416	Improved hen egg w
27	7	3.3	19	AAW74938	Human secreted pro
28	7	3.3	21	AAW80973	Sequence of modifi
29	7	3.3	21	AAW92065	N-terminal region
30	7	3.3	21	AAW05417	Signal peptide:
31	7	3.3	22	AAW00284	Human secreted pro
32	7	3.3	43	AAW57883	Arabidopsis thalia
33	7	3.3	44	AAW14432	Human secreted pro
34	7	3.3	46	AAW22521	Novel human colon
35	7	3.3	46	AAW02476	Human digestive sy
36	7	3.3	47	AAW02178	Human polypeptide
37	7	3.3	48	ABW5983	Human testicular a
38	7	3.3	48	AAW5286	Human reproductive
39	7	3.3	52	AAW00502	Human secreted pro
40	7	3.3	55	AAW19543	Amino acid sequenc
41	7	3.3	57	ABW41022	Peptide #8528 enco
42	7	3.3	57	ABW5106	Protein #7105 enco
43	7	3.3	57	AAW1879	Human brain expres
44	7	3.3	57	AAW4681	Human bone marrow
45	7	3.3	57	AAW20378	Peptide #6812 enco

ALIGNMENTS

RESULT 1
AAW40006
ID AAW40006 standard; peptide: 29 AA.

AC AAW40006;

DT 18-JUN-1998 (first entry)

Peptide effecting G-protein-coupled receptor activity.

G-protein-coupled receptor; GPCR: transmembrane domain; oligomerisation;
therapeutic composition; GPCR function; receptor monomeric form;
multimeric form; inhibition; GPCR-mediated process; GPCR binding;
treatment; disease; platelet activating factor.

OS Synthetic.
OS Homo sapiens.
XX
XX W09800538-A2.
PN

XX 08-JAN-1998.
PD
XX
XX 01-JUL-1997; 97WO-IB00814.
PF
XX
XX 01-JUL-1996; 96US-0021031.
PR
XX
XX (BIOS-) BIOSIGNAL INC.
PA (UYMO-) UNIV MONTREAL.
XX
XX
PI Bouvier M, Dennis M, Hebert TE;
XX
XX WPI; 1998-086964/08.
DR
XX
XX Peptide(s) or peptide leads affecting G protein-coupled receptor
PT activity - by altering receptor oligomerisation, useful in, e.g.
PR selectively modulating receptor function and treating neurological
PT or genetic diseases
XX
XX
PS Claim 20; Page 62; 75pp; English.
XX
XX Peptides AAM40002-08 are modelled on transmembrane domains one to seven,
CC respectively, of human platelet activating factor, a G-protein-coupled
CC receptor (GPCR) (sic) whose activity is affected by the formation of
CC oligomers. GPCRs have a recurring pattern unique for the transmembrane
CC domains. The peptides are characterised by the ability to selectively
CC affect oligomerisation of the GPCR from which it was designed. The
CC peptides may be used with suitable carriers in the preparation of
CC therapeutic compositions. They may be administered to selectively
CC modulate GPCR function by affecting the ratio of receptor monomeric
CC to multimeric forms. The compositions can be administered to inhibit
CC GPCR-mediated processes by modulating GPCR binding to its agonist or
CC antagonist. The peptides may also be used to prevent or treat diseases
CC involving GPCRs. They are also useful for in vitro and in vivo studies
CC of GPCRs.
XX
XX
SQ Sequence 29 AA;
XX
XX
XX Alignment Scores:
Pred. No.: 32.4 Length: 29
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 19 Gaps: 0
XX
XX US-09-835-992A-20 (1-687) x AAM40006 (1-29)
OY 103 TCTTTTTCCTAGTTTCTAATA 80
DB 11 SerpPhePheValPheLeuIle 18
XX
XX
XX RESULT 2
ID AAO11927
XX AAO11927 standard; Protein; 104 AA.
XX
XX AAO11927;
XX
XX 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 25819.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
PD
XX

PF 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI; 2001-514838/56.
DR N-PSDB; AA191858.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX
PS Claim 20; SEQ ID NO 25819; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activating/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
XX
XX
SQ Sequence 104 AA;
XX
XX
XX Alignment Scores:
Pred. No.: 28.8 Length: 104
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0
XX
XX US-09-835-992A-20 (1-687) x AAO11927 (1-104)
OY 351 TTGGGAGAGAAATGCTGTCG 374
DB 15 PheGlyLysLysIleLeuPheLeu 22
XX
XX
XX RESULT 3
ID ABB31093
XX ABB31093 standard; Peptide; 159 AA.
XX
XX ABB31093;
XX
XX 01-FEB-2002 (first entry)
DE Peptide #3744 encoded by breast cell single exon nucleic acid probe.
XX
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00662.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
PT Now spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX
PS Claim 27; SEQ ID NO 14061; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 159 AA;
SQ

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
Gaps: 0
DB: 22

US-09-835-992A-20 (1-687) x ABB31093 (1-159)
QY 479 CTGTCGCATTGTCCTTACTT 502
AC |
XX |
DB 120 LeuLeuAlaLeuCySPheLeuLeu 127
RESULT 4
ABB36288
ID ABB36288 standard; Peptide; 159 AA.
XX
AC ABB36288;
XX
DT 04-FEB-2002 (first entry)
XX
XX Peptide #3794 encoded by human foetal liver single exon probe.
DE
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
KM
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-493447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
XX
PS Claim 27; SEQ ID NO 28923; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC foetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 159 AA;
SQ

Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
Gaps: 0
DB: 22

US-09-835-992A-20 (1-687) x ABB36288 (1-159)
QY 479 CTGTCGCATTGTCCTTACTT 502
AC |
XX |
DB 120 LeuLeuAlaLeuCySPheLeuLeu 127
RESULT 5
ABB21653
ID ABB21653 standard; Protein; 159 AA.
XX
AC ABB21653;
XX
DT 23-JAN-2002 (first entry)
XX
XX Protein #3652 encoded by probe for measuring heart cell gene expression.
DE
XX
XX Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

```
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 15; SEQ ID NO 23423; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX AB21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX by measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosticating diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX Sequence 159 AA:
SQ
Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0
US-09-835-992a-20 (1-687) x ABB21653 (1-159)
QY 479 CTGTTGGCATTGTGTTCTTACTT 502
DB 120 LeuLeuAlaLeuCySpheLeuLeu 127
RESULT 6
AAM57060
ID AAM57060 standard; Protein: 159 AA.
AC AAM57060;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 29165.
XX
XX Human; brain expressed exon; gene expression analysis; probe:
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
```

```
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 29165; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 159 AA:
SQ
Alignment Scores:
Pred. No.: 27.7 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 22 Gaps: 0
US-09-835-992a-20 (1-687) x AAM57060 (1-159)
QY 479 CTGTTGGCATTGTGTTCTTACTT 502
DB 120 LeuLeuAlaLeuCySpheLeuLeu 127
RESULT 7
AAM69450
ID AAM69450 standard; Protein: 159 AA.
AC AAM69450;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 29756.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe:
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 29756; 658pp + Sequence Listing; English.
XX
XX
```

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

XX Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992a-20 (1-687) x AAM69450 (1-159)

OY 479 CTGTTGGCATGTGTTCTTACTT 502

DB 120 LeuLeuAlaLeuCySpHeLeuLeu 127

RESULT 8

AAM17284 standard; Protein: 159 AA.

XX AAM17284;

XX 12-OCT-2001 (first entry)

XX Peptide #3718 encoded by probe for measuring cervical gene expression.

XX Probe: human: microarray; gene expression; cervical epithelial cell;

XX cervical cancer.

XX Homo sapiens.

XX MO200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 27: SEQ ID NO 22110; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992a-20 (1-687) x AAM17284 (1-159)

OY 479 CTGTTGGCATGTGTTCTTACTT 502

DB 120 LeuLeuAlaLeuCySpHeLeuLeu 127

RESULT 9

AAM29784 standard; Protein: 159 AA.

XX AAM29784;

XX 17-OCT-2001 (first entry)

XX Peptide #3821 encoded by probe for measuring placental gene expression.

XX Probe: microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX MO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 27: SEQ ID NO 30053; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP:

XX see AAI131315-AA157946). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX CC predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders.

XX Sequence 159 AA;

Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 3.77% Indels: 0
 DB: 22 Gaps: 0

US-09-835-992a-20 (1-687) x AAM29784 (1-159)

QY 479 CTGTGGCATTGTTCTTACTT 502
 |||
 Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 10
 AAM04971
 ID AAM04971 standard; Protein; 159 AA.
 AC AAM04971;
 DT 09-OCT-2001 (first entry)

DE Peptide #3653 encoded by probe for measuring breast gene expression.
 XX
 KW Probe; human; breast disease; breast cancer; development disorder;
 KW Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200157270-A2.
 PD 09-AUG-2001.
 PF 29-JAN-2001; 2001WO-US00661.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 PS
 XX
 PS Claim 27; SEQ ID NO 13711; 322pp; English.

CC The present invention relates to novel single exon nucleic acid probes
 CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for measuring human gene expression in
 CC a human breast sample, where the probe hybridises at high stringency to a
 CC nucleic acid expressed in the human breast. The probes are useful for
 CC predicting, diagnosing, grading, staging, monitoring and prognosing
 CC diseases of the human breast, particularly those diseases with polygenic
 CC aetiology. The diseases include: breast cancer, disorders of development,
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative
 CC breast disease and non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 159 AA;

Alignment Scores: 27.7 Length: 159
 Pred. No.: 8.00 Matches: 8
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 3.77% Gaps: 0

US-09-835-992a-20 (1-687) x AAM04971 (1-159)

QY 479 CTGTGGCATTGTTCTTACTT 502
 |||
 Db 120 LeuLeuAlaLeuCySpheLeuLeu 127

RESULT 11
 ABG39070
 ID ABG39070 standard; Peptide; 159 AA.
 AC ABG39070;
 DT 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 28735.
 XX
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenar syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 PD 15-NOV-2001.
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS
 XX
 PS Claim 27; SEQ ID NO 28735; 634pp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exon should be assigned to a single gene; a peptide comprising one
 CC of 1201 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC hemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 159 AA;

CC Alignment Scores:

Pred. No.:	27.7	Length:	159
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	23	Gaps:	0

US-09-835-992a-20 (1-687) x ABC93070 (1-159)

QY 479 CTGTTGGATTGTTCTTACTT 502

Db 120 LeuLeuAlaLeuLeuLeuLeu 127

RESULT 12

ABD09061 standard; Protein: 275 AA.

AC ABD09061;

DT 26-JUN-2002 (first entry)

DE Thermus caldophilus GK24 lactose permease SEQ ID NO:13.

KM Thermus caldophilus GK24; heat resistant enzyme; cosmetic; food;
 KM DNA polymerase; equalysin I; beta-galactosidase; lactose permease.

OS Thermus caldophilus.

PN KR2001019888-A.

PD 15-MAR-2001.

PF 31-AUG-1999; 99KR-0036564.

PR 31-AUG-1999; 99KR-0036564.

PA (KORE-) KOREA RES INST BIOSCIENCE & BIOTECHNOLOG.

PI Cho SJ, Jung BW, Kim HJ, Kim JS, Kim MG, Kim SY, Kim YS, Ko JH;
 PI Ko SH, Lee DS, Lee JS, Park JA, Park JH, Shin HJ;

DR WPI; 2001-512441/56.

PT Base sequence and amino acid sequence of cosmid clone 3 isolated from
 PT thermus caldophilus -

PS Example 2; Page 109-111; 130pp; Korean.

CC The present invention describes the nucleotide sequence of the cosmid

CC clone 3 isolated from Thermus caldophilus GK24, as given in the present
 CC invention. Proteins encoded by the cosmid clone 3 are heat resistant
 CC enzymes which can be used in the fields of cosmetic and food industries.
 CC Thermus caldophilus GK24 has the optimal growth temperature of 72 plus
 CC degrees Celsius and produces heat resistant enzymes including DNA
 CC polymerase, equalysin I, and beta-galactosidase. The DNA polymerase can
 CC effectively be used in PCR which is performed at high temperature. The
 CC present sequence represents lactose permease encoded by the Thermus
 CC caldophilus GK24 cosmid clone 3.

CC Sequence 275 AA;

CC Alignment Scores:

Pred. No.:	26.3	Length:	275
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	22	Gaps:	0

US-09-835-992a-20 (1-687) x ABD09061 (1-275)

QY 594 GTCTTGGATTGTTACTTCTA 617

Db 196 ValLeuTrpIleLeuLeuPheLeu 203

RESULT 13

AAR25835 standard; Protein: 342 AA.

AC AAR25835;

DT 21-JAN-1993 (first entry)

DE Human platelet activating factor receptor.

KM PAF; coagulation; blood clotting; clone pPAF.

OS Homo sapiens.

PN WO9212244-A.

PD 23-JUL-1992.

PF 27-DEC-1991; 91WO-JP01779.

PR 27-DEC-1990; 90JP-0407119.

PA (NISR) JAPAN TOBACCO INC.

PI Honda Z, Nakamura M, Shimizu T;

DR WPI; 1992-268667/32.

DR N-PSDB; AAQ26960.

PT DNA coding for human or guinea pig platelet activating factor
 PT receptor - for prodn. of recombinant PAF receptor for screening
 PT potential PAF agonists-antagonists

PS Claim 3; Fig 5; 50pp; Japanese.

PI The cDNA coding for human PAF receptor was isolated in a 1780 bp
 PI clone designated pPAF. The clone consisted of a 112bp 5'-UTR (see
 PI AAQ26961), a 1029bp ORF (AAQ26960) and a 639bp 3'-UTR (see
 PI AAQ26962).

CC The cDNA can be used to transform a suitable host cell line (e.g. COS
 CC cells) for expression of the PAF receptor protein. The receptor
 CC itself may be useful in screening of potential PAF agonists/
 CC antagonists and in assays for PAF. See also AAQ26957-026959.

PS Sequence 342 AA;

CC Alignment Scores:

Pred. No.: 25.8 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 13 Gaps: 0

US-09-835-992A-20 (1-687) x AAR25835 (1-342)

OY 103 TCTTTTTCCTAGTTTTCATA 80
 DB 194 SerPhePheLeuValPheLeuile 201

RESULT 14

AAV49550 standard; Protein; 342 AA.

AAV49550; (first entry)

Human platelet activating factor receptor protein sequence.

Human: coding sequence polymorphism; vascular pathology gene;

polymorphic site; phenotype correlation; forensic; paternity testing;

medicine; genetic analysis; vascular disease.

Homo sapiens.

WO950454-A2.

07-OCT-1999.

26-MAR-1999; 99WO-US06473.

01-APR-1998; 98US-0054272.

(WHED) WHITEHEAD INST BIOMEDICAL RES.

Lander ES, Daley GQ, Cargill M, Ireland JS, Rozen SG;

MP1; 1999-620066/53.

N-PSDB; AA32159.

Determination of polymorphisms in genes, especially those identifying

predisposition to vascular disease

Disclosure: Fig 3; 134pp; English.

AA32159 to AA32194 represent reference alleles for specifically
 claimed nucleic acid sequences from the present invention which comprise
 CC polymorphic sites as given in a table in the specification, selected
 from 92 single nucleotide polymorphisms in which the nucleotide at the
 CC polymorphic site is different from a nucleotide at the same site in a
 reference allele. The nucleic acids, and primers and probes, are used to
 identify polymorphisms, which may predispose an individual to disease,
 CC especially a vascular disease. They can also be used in phenotype
 CC correlations, forensics, paternity testing, medicine or genetic
 CC analysis. AAV49550 to AAV49573 represent the proteins which correspond
 to some of the reference alleles.

Sequence 342 AA:

Alignment Scores:

Pred. No.: 25.8 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 20 Gaps: 0

US-09-835-992A-20 (1-687) x AAV49550 (1-342)

OY 103 TCTTTTTCCTAGTTTTCATA 80
 DB 194 SerPhePheLeuValPheLeuile 201

RESULT 15

ABB56381 standard; Protein; 342 AA.

ABB56381;

18-FEB-2002 (first entry)

Non-endogenous human GPCR protein, SEQ ID NO: 555.

Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
 constitutively activated GPCR; agonist; disease.

Homo sapiens.

Synthetic.

WO20017172-A2.

18-OCT-2001.

05-APR-2001; 2001WO-US11098.

07-APR-2000; 2000US-195747P.

(AREN-) ARENA PHARM INC.

Lehmann-Bruinsma K, Liaw CW, Lin I;

MP1; 2001-648759/74.

N-PSDB; ABI98017.

Identifying agonists of G protein-coupled receptors (GPCRs) for use in
 disease treatment, comprises contacting candidate compounds with

versions of GPCRs

Claim 1; Page 356-357; 394pp; English.

The invention relates to G protein-coupled receptors (GPCRs) for which
 the endogenous ligand has been identified. Non-endogenous
 CC constitutively activated versions of known GPCRs are used in the
 CC invention for the direct identification of candidate compounds as
 CC receptor agonists, inverse agonists or partial agonists. Such
 CC agonists are useful as therapeutic agents for diseases or disorders
 CC associated with GPCRs. The present sequence is a non-endogenous
 CC version of a known human GPCR.

Sequence 342 AA:

Alignment Scores:
 Pred. No.: 25.8 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 22 Gaps: 0

US-09-835-992A-20 (1-687) x ABB56381 (1-342)

OY 103 TCTTTTTCCTAGTTTTCATA 80
 DB 194 SerPhePheLeuValPheLeuile 201

Search completed: January 14, 2003, 17:27:19
 job time : 35.8408 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 ; Search time 9.44148 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992a-20

Perfect score: 212
Sequence: 1 attcaatttcattcaaac.....cttgatcagcacagtgtt 687

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0	
Delop 6.0 , Delext 7.0	

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+ n2p.model -DEV-x1p
-O/cgn2.1/USPTO.spool/US09835992/runal.14012003.161527.4050/app.query.fasta.1.3932
-DB-issued_Patents_AA -QFWT-fastan -SUFFIX-olig.ra1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.cdi
-LIST=45 -DOCALLIN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFWT=pio -NORM-ext -HEAVY=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992 -ECGN.1.1.49 -runat.14012003.161527.4050 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database:

- 1: /cgn2.6/prodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2.6/prodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2.6/prodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2.6/prodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2.6/prodata/1/1aa/PCTUS.COMB.pep.*
- 6: /cgn2.6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	542	1 US-08-246-583-3	Sequence 3, Appl1
2	7	3.3	18	6 5182195-2	Patent No. 5182195
3	7	3.3	19	4 US-09-149-476-534	Sequence 534, App
4	7	3.3	21	6 5182195-7	Patent No. 5182195
5	7	3.3	44	4 US-09-288-143-88	Sequence 88, Appl1
6	7	3.3	62	1 US-08-488-861-6	Sequence 6, Appl1
7	7	3.3	62	4 US-08-973-297-6	Sequence 6, Appl1
8	7	3.3	62	5 PCT-US96-06511-6	Sequence 3466, App
9	7	3.3	71	4 US-09-134-001C-3466	Sequence 417, Appl
10	7	3.3	187	4 US-09-149-476-417	Sequence 12, Appl
11	7	3.3	222	1 US-07-869-933-12	Sequence 28, Appl1
12	7	3.3	222	1 US-07-869-933-28	Sequence 12, Appl1

13	7	3.3	222	4 US-09-103-663-12	Sequence 12, Appl1
14	7	3.3	222	4 US-09-103-663-28	Sequence 28, Appl1
15	7	3.3	267	2 US-08-557-128-4	Sequence 4, Appl1
16	7	3.3	267	4 US-09-242-690A-36	Sequence 36, Appl1
17	7	3.3	326	1 US-07-603-133B-25	Sequence 25, Appl1
18	7	3.3	380	4 US-08-857-076-110	Sequence 110, App
19	7	3.3	427	4 US-09-134-001C-5143	Sequence 5143, App
20	7	3.3	527	1 US-08-246-583-2	Sequence 2, Appl1
21	7	3.3	538	2 US-08-933-227-5	Sequence 5, Appl1
22	7	3.3	554	4 US-09-321-276-4	Sequence 4, Appl1
23	7	3.3	554	4 US-08-916-481-2	Sequence 2, Appl1
24	7	3.3	563	4 US-08-916-481-3	Sequence 3, Appl1
25	7	3.3	663	4 US-09-196-293-5	Sequence 5, Appl1
26	7	3.3	663	4 US-08-209-603E-5	Sequence 5, Appl1
27	7	3.3	663	4 US-08-235-836C-70	Sequence 70, Appl
28	7	3.3	663	4 US-08-235-836C-78	Sequence 78, Appl
29	7	3.3	681	4 US-08-760-615-4	Sequence 4, Appl1
30	7	3.3	681	4 US-08-760-615-6	Sequence 6, Appl1
31	7	3.3	700	4 US-08-235-836C-74	Sequence 74, Appl
32	7	3.3	1698	4 US-09-315-793-12	Sequence 12, Appl
33	7	3.3	1724	4 US-08-857-076-12	Sequence 12, Appl
34	7	3.3	1956	4 US-08-843-417-2	Sequence 2, Appl1
35	7	3.3	1957	4 US-08-669-656A-2	Sequence 2, Appl1
36	7	3.3	1957	4 US-08-669-656A-8	Sequence 8, Appl1
37	7	3.3	2132	4 US-08-197-793-3	Sequence 3, Appl1
38	6	2.8	10	1 US-08-636-176-3	Sequence 3, Appl1
39	6	2.8	10	2 US-08-197-793-3	Sequence 3, Appl1
40	6	2.8	10	5 PCT-US95-01618-3	Sequence 3, Appl1
41	6	2.8	11	1 US-08-338-634-4	Sequence 4, Appl1
42	6	2.8	11	1 US-08-787-347-67	Sequence 67, Appl1
43	6	2.8	11	5 PCT-US95-16415-1	Sequence 1, Appl1
44	6	2.8	14	1 US-08-232-453A-40	Sequence 40, Appl1
45	6	2.8	15	4 US-09-027-900-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-246-583-3
Sequence 3, Application US/08246583
Patent No. 5750394
GENERAL INFORMATION:
APPLICANT: Palese, Peter
TITLE OF INVENTION: IDENTIFICATION AND USE OF ANTIVIRAL
COMPOUNDS THAT INHIBIT INTERACTION OF HOST CELL PROTEINS
TITLE OF INVENTION: AND VIRAL PROTEINS REQUIRED FOR VIRAL REPLICATION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 AVENUE OF THE AMERICAS
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/246,583
FILING DATE: 20-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 6923-040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 542 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-246-583-3

Alignment Scores:
Pred. No.: 4.97 Length: 542
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-20 (1-687) x US-08-246-583-3 (1-542)

QY 497 TTAAGTCTGTCCTCCAGGAAAC 520
|||||
Db 349 LeuLeuSerProlySgluasn 356

RESULT 2
5182195-2

Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO, KAISHO, YOSHIHIKO, YOSHIMURA, KOJI
DEFICIENT YEASTS
NUMBER OF SEQUENCES: 71
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,140
FILING DATE: 09-NOV-1988
SEQ ID NO: 2
LENGTH: 18
5182195-2

Alignment Scores:
Pred. No.: 80.7 Length: 18
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 6 Gaps: 0

US-09-835-992A-20 (1-687) x 5182195-2 (1-18)

QY 479 CTTGTCATGCTGTTCTTA 499
|||||
Db 6 LeuLeuAlaLeuCysPheLeu 12

RESULT 3
US-09-149-476-534
Sequence 534, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
EARLIER FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/056, 886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057, 761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047, 595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043, 578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047, 501

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043, 670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056, 632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048, 964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057, 650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056, 884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057, 669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049, 610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061, 060
EARLIER FILING DATE: 1997-10-02

Alignment Scores:

Pred. No.: 80.1 Length: 19
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 3.30% Indels: 0
IDB: 4 Gaps: 0

US-09-835-992a-20 (1-687) x US-09-149-476-534 (1-19)

QY 606 TTACTGTCCTAATTATTC 626
Db 4 Leuleupheuleuleuphe 10

RESULT 4

5182195-7

Patent No. 5182195
APPLICANT: NAKAHAMA, KAZUO; KAISHO, YOSHIMITSU, YOSHIMURA, KOJI
TITLE OF INVENTION: METHOD FOR INCREASING USING PROTEASE

DEFICIENT YEASTS

NUMBER OF SEQUENCES: 71

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/269,140

FILING DATE: 09-NOV-1988

SEQ ID NO: 7:

LENGTH: 21

5182195-7

Alignment Scores:

Pred. No.: 79 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 3.30% Indels: 0
DB: 6 Gaps: 0

US-09-835-992a-20 (1-687) x 5182195-7 (1-21)

QY 479 CTGTTGGCATTGTGTTCTTA 499

|||||
Db 6 leuLeuAlaLeucyspheleu 12

RESULT 5
US-09-288-143-88
; Sequence 88, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/US98/21142
; EARLIER FILING DATE: 1998-10-08
; EARLIER APPLICATION NUMBER: 60/061,463
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,529
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/071,498
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,527
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,536
; EARLIER FILING DATE: 1997-10-09
; EARLIER APPLICATION NUMBER: 60/061,532
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals stop translation
US-09-288-143-88

Alignment Scores:
Pred. No.: 71.1 Length: 44
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
Gaps: 0
DB: 4

US-09-835-992A-20 (1-687) x US-09-288-143-88 (1-44)

QY 603 ATTTCAGTCTTCCTATTTTA 623
|||||
Db 5 lleLeuAlaLeucyspheleu 11

RESULT 6
US-08-488-961-6
; Sequence 6, Application US/08488961
; Patent No. 5606042
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel S.
; APPLICANT: Walker, John C.
; TITLE OF INVENTION: Glycine and Phaseolus
; TITLE OF INVENTION: alpha-D-Galactosidases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Reising, Ethington, Barnard & Perry
; STREET: P. O. Box 4390
; CITY: Troy
; STATE: Michigan
; COUNTRY: US
; ZIP: 48099-4390
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,961
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: P-320 (UMO)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 689-3500
; TELEFAX: (810) 689-4071
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-488-961-6

Alignment Scores:
Pred. No.: 67.7 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
Gaps: 0
DB: 1

US-09-835-992A-20 (1-687) x US-08-488-961-6 (1-62)

QY 482 TTGGCAGTGTCTTCTACTT 502
|||||
Db 20 leuAlaLeucyspheleu 26

RESULT 7
US-08-973-297-6
; Sequence 6, Application US/08973297
; Patent No. 6184017
; GENERAL INFORMATION:
; APPLICANT: Smith, Daniel S.
; APPLICANT: Walker, John C.
; TITLE OF INVENTION: Glycine and Phaseolus
; TITLE OF INVENTION: alpha-D-Galactosidases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 NO. 6184017thwestern Hwy., Suite 410
; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: US
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,297
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Kohn, Kenneth I.
; REGISTRATION NUMBER: 30,955
; REFERENCE/DOCKET NUMBER: 0994.00050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 539-5050
; TELEFAX: (810) 539-5055
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

```

: TOPOLOGY: Linear
US-08-973-297-6

Alignment Scores:
Pred. No.: 67.7
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
DB: 4

US-09-835-992a-20 (1-687) x US-08-973-297-6 (1-62)
OY 482 TTGCATTGTTCTTACTT 502
DB 20 LeuAlaLeuCySpheLeuLeu 26

RESULT 8
PCT-US96-06511-6
: Sequence 6, Application PC/TUS9606511
: GENERAL INFORMATION:
: APPLICANT: Smith, Daniel S.
: TITLE OF INVENTION: Glycine and Phascolus
: TITLE OF INVENTION: alpha-D-Galactosidases
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kohn & Associates
: STREET: 30500 Northwestern Hwy., Suite 410
: CITY: Farmington Hills
: STATE: Michigan
: COUNTRY: US
: ZIP: 48334
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US96/06511
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Kohn, Kenneth I.
: REGISTRATION NUMBER: 30,955
: REFERENCE/DOCKET NUMBER: 0994.00050
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (810) 539-5050
: TELEFAX: (810) 539-5055
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 62 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: PCT-US96-06511-6

Alignment Scores:
Pred. No.: 67.7
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
DB: 5

US-09-835-992a-20 (1-687) x PCT-US96-06511-6 (1-62)
OY 482 TTGCATTGTTCTTACTT 502
DB 20 LeuAlaLeuCySpheLeuLeu 26

RESULT 9
US-09-134-001C-3466

```

```

: Sequence 3466, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 3466
: LENGTH: 71
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3466

Alignment Scores:
Pred. No.: 66.4
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.32%
DB: 4

US-09-835-992a-20 (1-687) x US-09-134-001C-3466 (1-71)
OY 240 CCATCATTAGATTAATGAA 220
DB 13 ProseleuArgileleGlu 19

RESULT 10
US-09-149-476-417
: Sequence 417, Application US/09149476
: Patent No. 6420526
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: 186 Human Secreted proteins
: FILE REFERENCE: P2002P1
: CURRENT APPLICATION NUMBER: US/09/149,476
: CURRENT FILING DATE: 1998-09-08
: EARLIER APPLICATION NUMBER: PCT/US98/04493
: EARLIER FILING DATE: 1998-03-06
: EARLIER APPLICATION NUMBER: 60/040,162
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,333
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/038,621
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,626
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,334
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,336
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/040,163
: EARLIER FILING DATE: 1997-03-07
: EARLIER APPLICATION NUMBER: 60/047,600
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,615
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,597
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,502
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,633
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,583
: EARLIER FILING DATE: 1997-05-23
: EARLIER APPLICATION NUMBER: 60/047,617

```

1	EARLIER	FILING DATE:	1997-05-23
2	EARLIER	APPLICATION NUMBER:	60/047, 6118
3	EARLIER	FILING DATE:	1997-05-23
4	EARLIER	APPLICATION NUMBER:	60/047, 5038
5	EARLIER	FILING DATE:	1997-05-23
6	EARLIER	APPLICATION NUMBER:	60/047, 5922
7	EARLIER	FILING DATE:	1997-05-23
8	EARLIER	APPLICATION NUMBER:	60/047, 5818
9	EARLIER	FILING DATE:	1997-05-23
10	EARLIER	APPLICATION NUMBER:	60/047, 5848
11	EARLIER	FILING DATE:	1997-05-23
12	EARLIER	APPLICATION NUMBER:	60/047, 5000
13	EARLIER	FILING DATE:	1997-05-23
14	EARLIER	APPLICATION NUMBER:	60/047, 5988
15	EARLIER	FILING DATE:	1997-05-23
16	EARLIER	APPLICATION NUMBER:	60/047, 6113
17	EARLIER	FILING DATE:	1997-05-23
18	EARLIER	APPLICATION NUMBER:	60/047, 5822
19	EARLIER	FILING DATE:	1997-05-23
20	EARLIER	APPLICATION NUMBER:	60/047, 5968
21	EARLIER	FILING DATE:	1997-05-23
22	EARLIER	APPLICATION NUMBER:	60/047, 6122
23	EARLIER	FILING DATE:	1997-05-23
24	EARLIER	APPLICATION NUMBER:	60/047, 6322
25	EARLIER	FILING DATE:	1997-05-23
26	EARLIER	APPLICATION NUMBER:	60/047, 6001
27	EARLIER	FILING DATE:	1997-05-23
28	EARLIER	APPLICATION NUMBER:	60/043, 5800
29	EARLIER	FILING DATE:	1997-04-11
30	EARLIER	APPLICATION NUMBER:	60/043, 5668
31	EARLIER	FILING DATE:	1997-04-11
32	EARLIER	APPLICATION NUMBER:	60/043, 3144
33	EARLIER	FILING DATE:	1997-04-11
34	EARLIER	APPLICATION NUMBER:	60/043, 5656
35	EARLIER	FILING DATE:	1997-04-11
36	EARLIER	APPLICATION NUMBER:	60/043, 3111
37	EARLIER	FILING DATE:	1997-04-11
38	EARLIER	APPLICATION NUMBER:	60/043, 6711
39	EARLIER	FILING DATE:	1997-04-11
40	EARLIER	APPLICATION NUMBER:	60/043, 6744
41	EARLIER	FILING DATE:	1997-04-11
42	EARLIER	APPLICATION NUMBER:	60/043, 6656
43	EARLIER	FILING DATE:	1997-04-11
44	EARLIER	APPLICATION NUMBER:	60/043, 3122
45	EARLIER	FILING DATE:	1997-04-11
46	EARLIER	APPLICATION NUMBER:	60/043, 3133
47	EARLIER	FILING DATE:	1997-04-11
48	EARLIER	APPLICATION NUMBER:	60/043, 6722
49	EARLIER	FILING DATE:	1997-04-11
50	EARLIER	APPLICATION NUMBER:	60/043, 3155
51	EARLIER	FILING DATE:	1997-04-11
52	EARLIER	APPLICATION NUMBER:	60/048, 9744
53	EARLIER	FILING DATE:	1997-06-06
54	EARLIER	APPLICATION NUMBER:	60/056, 8866
55	EARLIER	FILING DATE:	1997-08-22
56	EARLIER	APPLICATION NUMBER:	60/056, 8777
57	EARLIER	FILING DATE:	1997-08-22
58	EARLIER	APPLICATION NUMBER:	60/056, 8899
59	EARLIER	FILING DATE:	1997-08-22
60	EARLIER	APPLICATION NUMBER:	60/056, 8933
61	EARLIER	FILING DATE:	1997-08-22
62	EARLIER	APPLICATION NUMBER:	60/056, 6300
63	EARLIER	FILING DATE:	1997-08-22
64	EARLIER	APPLICATION NUMBER:	60/056, 8788
65	EARLIER	FILING DATE:	1997-08-22
66	EARLIER	APPLICATION NUMBER:	60/056, 6622
67	EARLIER	FILING DATE:	1997-08-22
68	EARLIER	APPLICATION NUMBER:	60/056, 8722
69	EARLIER	FILING DATE:	1997-08-22

EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,550
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 13

US-09-103-663-12

Sequence 12, Application US/09103663D

Patent No. 6171803

GENERAL INFORMATION:

APPLICANT: Kinet et al.

TITLE OF INVENTION: Isolation, characterization, and use of the human beta

TITLE OF INVENTION: subunit of the high affinity receptor for

TITLE OF INVENTION: immunoglobulin E.

FILE REFERENCE: 50490

CURRENT APPLICATION NUMBER: US/09/103,663D

EARLIER FILING DATE: 1998-06-23

EARLIER FILING DATE: 1992-04-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 222

TYPE: PRT

ORGANISM: Rattus sp.

US-09-103-663-12

Alignment Scores:

Pred. No.: 56.5 Length: 222

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.30% Indels: 0

DB: 4 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-103-663-12 (1-222)

OY 69 AATAACATTAGTATTGAAAA 89

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 14

US-09-103-663-28

Sequence 28, Application US/09103663D

Patent No. 6171803

GENERAL INFORMATION:

APPLICANT: Kinet et al.

TITLE OF INVENTION: Isolation, characterization, and use of the human beta

TITLE OF INVENTION: subunit of the high affinity receptor for

TITLE OF INVENTION: immunoglobulin E.

FILE REFERENCE: 50490

CURRENT APPLICATION NUMBER: US/09/103,663D

EARLIER FILING DATE: 1998-06-23

EARLIER FILING DATE: 1992-04-16

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 28

LENGTH: 222

TYPE: PRT

ORGANISM: Homo sapiens

US-09-103-663-28

Alignment Scores:

Pred. No.: 56.5 Length: 222

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.30% Indels: 0

DB: 4 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-103-663-28 (1-222)

OY 69 AATAACATTAGTATTGAAAA 89

Db 134 AsnAsnIleSerIleArgLys 140

RESULT 15

US-08-557-128-4

Sequence 4, Application US/08557128

Patent No. 5849524

GENERAL INFORMATION:

APPLICANT: KONDO, Keiji

APPLICANT: KAJIMURA, Susumu

TITLE OF INVENTION: TRANSFORMATION SYSTEMS FOR THE YEAST

TITLE OF INVENTION: CANDIDA UTILIS AND THE EXPRESSION OF HETEROLOGOUS GENES

TITLE OF INVENTION: THEREWITH

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/557,128

FILING DATE: 25-JAN-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/JP95/01005

FILING DATE: 25-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-129287

FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-285823

FILING DATE: 26-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-135015

FILING DATE: 25-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 49441/108

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-557-128-4

Alignment Scores:

Pred. No.: 55 Length: 267

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.32% Indels: 0

DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x US-08-557-128-4 (1-267)

OY 507 GAACATAAGTAAGAACACAAAT 487

Db 80 GluIleuSerIlyslYshIsAsn 86

Wed Jan 15 08:41:06 2003

Search completed: January 14, 2003, 17:42:09
Job time : 10.6915 secs

us-09-835-992a-20.olg.ra1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 6.57721 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 212
Sequence: 1 atttaatttcatttaaac.....cttgatgacacagttgtt 687

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xip
-O/cgn2.1/USPTO.spool/US09835992/rumat.14012003.161528.4118/app.query.fasta.1.3932
-DB=Published.Applications_AA -QFMT=fastan -SUFFIX=olg.rapb -MINMATCH=0.1
-LOOPEC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAJLEN=2000000000 -USER=US09835992 -RCGN.1.1-28 -rumat.14012003.161528.4118
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEOBURY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published.Applications_AA:*

1:	/cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
2:	/cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
8:	/cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep:*
10:	/cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
12:	/cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
13:	/cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
14:	/cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3.8	159	10	US-09-864-761-36951	Sequence 36951, A
2	3.8	327	12	US-10-084-206-3	Sequence 3, Appl1
3	3.3	57	10	US-09-864-761-40404	Sequence 40404, A
4	3.3	81	9	US-10-002-344A-243	Sequence 243, App

5	7	3.3	127	10	US-09-764-877-1466	Sequence 1466, Ap
6	7	3.3	197	10	US-09-780-717-47	Sequence 47, Appl
7	7	3.3	241	10	US-09-740-668A-26	Sequence 26, Appl
8	7	3.3	241	12	US-10-059-964-62	Sequence 62, Appl
9	7	3.3	267	10	US-09-908-855-36	Sequence 36, Appl
10	7	3.3	334	10	US-09-815-242-13249	Sequence 13249, A
11	7	3.3	380	10	US-09-205-658-110	Sequence 110, App
12	7	3.3	380	10	US-09-844-353A-110	Sequence 110, App
13	7	3.3	481	10	US-09-815-242-10888	Sequence 10888, A
14	7	3.3	554	10	US-09-800-396-4	Sequence 4, Appl1
15	7	3.3	758	10	US-09-925-299-859	Sequence 859, App
16	7	3.3	1724	10	US-09-305-658-112	Sequence 12, Appl
17	7	3.3	1724	10	US-09-844-353A-112	Sequence 12, Appl
18	6	2.8	9	10	US-09-834-765-444	Sequence 444, App
19	6	2.8	10	8	US-08-452-843A-13	Sequence 13, Appl
20	6	2.8	10	10	US-09-834-765-82	Sequence 82, Appl
21	6	2.8	10	10	US-09-834-765-611	Sequence 611, App
22	6	2.8	17	9	US-09-909-460-67	Sequence 67, Appl
23	6	2.8	17	10	US-09-949-196-4	Sequence 4, Appl1
24	6	2.8	20	10	US-09-742-732-5	Sequence 5, Appl1
25	6	2.8	22	9	US-10-001-876-130	Sequence 32, Appl
26	6	2.8	22	10	US-09-949-196-32	Sequence 36, Appl
27	6	2.8	22	10	US-09-949-196-36	Sequence 40, Appl
28	6	2.8	22	10	US-09-949-196-40	Sequence 113, App
29	6	2.8	25	12	US-10-001-879-113	Sequence 39910, A
30	6	2.8	27	10	US-09-864-761-39910	Sequence 10, App
31	6	2.8	28	10	US-09-929-818-104	Sequence 327, App
32	6	2.8	30	10	US-09-071-838-227	Sequence 47089, A
33	6	2.8	31	10	US-09-864-761-38920	Sequence 47089, A
34	6	2.8	32	10	US-09-864-761-47089	Sequence 47089, A
35	6	2.8	34	10	US-09-864-761-34176	Sequence 47949, A
36	6	2.8	36	10	US-09-864-761-47949	Sequence 44137, A
37	6	2.8	37	10	US-09-864-761-44137	Sequence 42251, A
38	6	2.8	39	10	US-09-864-761-42251	Sequence 46929, A
39	6	2.8	40	10	US-09-864-761-46929	Sequence 44190, A
40	6	2.8	41	10	US-09-864-761-44190	Sequence 866, App
41	6	2.8	43	10	US-09-864-761-44190	Sequence 199, App
42	6	2.8	43	10	US-09-761-847-866	Sequence 147, App
43	6	2.8	48	12	US-10-001-879-199	
44	6	2.8	49	9	US-09-843-676-147	
45	6	2.8	49	9	US-09-766-253-147	

ALIGNMENTS

RESULT 1
US-09-864-761-36951
Sequence 36951, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36951
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007374.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: AA485663.1, EVALUATE 1.00e-67
US-09-835-992a-20 (1-687) x US-09-864-761-36951 (1-159)
Alignment Scores:
Pred. No.: 5.64 Length: 159
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 10 Gaps: 0
Oy 479 CTTGTCGATGTTCTTACTT 502
Db 120 LeuleuAlaLeuCyspheLeuleu 127
RESULT 2
US-10-084-206-3
; Sequence 3, Application US/10084206
; Patent No. US20020106741A1
; GENERAL INFORMATION:
; APPLICANT: Li et al.
; TITLE OF INVENTION: G-Protein Receptor HTNAD29
; FILE REFERENCE: PF191D1C1
; CURRENT APPLICATION NUMBER: US/10/084,206
; PRIOR APPLICATION NUMBER: US95/07288
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/468,534
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 09/399,095
; PRIOR FILING DATE: 1999-09-20
```

```

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Version 3.1
; SEQ ID NO 3
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-084-206-3
Alignment Scores:
Pred. No.: 4.97 Length: 327
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 12 Gaps: 0
US-09-835-992a-20 (1-687) x US-10-084-206-3 (1-327)
Oy 103 TCTTTTCTCAGTTTCTAATA 80
Db 190 SerPhepheValpheLeulle 197
RESULT 3
US-09-864-761-40404
; Sequence 40404, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 40404
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078591.18
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
OTHER INFORMATION: SWISSPROT HIT: Q26563, EVALUE 2.70e+00
OTHER INFORMATION: EST_HUMAN HIT: BE169919.1, EVALUE 6.20e+00
US-09-864-761-40404

Alignment Scores:
Pred. No.: 62.9
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
Length: 57
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-835-992a-20 (1-687) x US-09-864-761-40404 (1-57)

OY 626 CTGAACTCCATTTTCCCA 646

Db 15 LeuLysLeuHisPheSerPro 21

RESULT 4

US-10-002-344A-243
Sequence 243, Application US/10002344A
Patent No. US20020172939A1
GENERAL INFORMATION:
APPLICANT: Recipon, Heve
APPLICANT: Sun, Yongming
APPLICANT: Chen, Sol-yu
APPLICANT: Liu, Chenghua
APPLICANT: Turner, Leah
TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
FILE REFERENCE: DEX-0241
CURRENT APPLICATION NUMBER: US/10/002,344A
CURRENT FILING DATE: 2002-03-15
PRIOR APPLICATION NUMBER: US 60/242,998
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 277
SOFTWARE: Patentin version 3.1
SEQ ID NO 243
LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapien
US-10-002-344A-243

Alignment Scores:
Pred. No.: 59.1
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
Length: 81
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-835-992a-20 (1-687) x US-10-002-344A-243 (1-81)

OY 603 ATTTACTGTCTTATTTTA 623

Db 73 ILleuLeuPheLeuLeuLeu 79

RESULT 5

US-09-764-877-1466

Sequence 1466, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1466
LENGTH: 127
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1466

Alignment Scores:
Pred. No.: 54.6
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
Length: 127
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-835-992a-20 (1-687) x US-09-764-877-1466 (1-127)

OY 312 GCCTGTTCAGCTGATC 332

Db 80 AlacysValGlnAlaLeuLeu 86

RESULT 6
US-09-780-717-47
Sequence 47, Application US/09780717
Patent No. US20010044941A1
GENERAL INFORMATION:
APPLICANT: Helentjaris, Tim
APPLICANT: Bates, Nic
APPLICANT: Allen, Stephen M.
TITLE OF INVENTION: of use
FILE REFERENCE: 035718/208677
CURRENT APPLICATION NUMBER: US/09/780,717
CURRENT FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: US 60/181,509
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 197
TYPE: PRT
ORGANISM: Glycine max
US-09-780-717-47

Alignment Scores:
Pred. No.: 50.5
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
Length: 197
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-835-992a-20 (1-687) x US-09-780-717-47 (1-197)

OY 29 CAAAGCATTTTATCTAT 49

Db 45 GlnSerIleIleuSerTyr 51

RESULT 7

US-09-740-668A-26
; Sequence 26, Application US/09740668A
; Patent No. US20020076700A1
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; TITLE OF INVENTION: No. US20020076700A1 polypeptides and nucleic acids encoding same
; FILE REFERENCE: 15966-537 CIP
; CURRENT APPLICATION NUMBER: US/09/740, 668A
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: PCT/US99/29584
; PRIOR FILING DATE: 1999-12-17
; PRIOR APPLICATION NUMBER: 09/465, 512
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113, 485
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/112, 837
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)..(241)
; OTHER INFORMATION: wherein Xaa is any amino acid as defined in the specification.
US-09-740-668A-26

Alignment Scores:
Pred. No.: 48.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-740-668A-26 (1-241)

OY 324 GCTCTGATCATATTCCTTTA 344
|||||
Db 126 AAlaulllellepheleuenu 132

RESULT 8
US-10-059-964-62
; Sequence 62, Application US/10059964
; Patent No. US20020120115A1
; GENERAL INFORMATION:
; APPLICANT: Rikihisa, Yasuko
; APPLICANT: Ohashi, No. US20020120115A1
; TITLE OF INVENTION: Outer Membrane Protein of Ehrlichia Canis and Ehrlichia
; FILE REFERENCE: 22727/04021
; CURRENT APPLICATION NUMBER: US/10/059, 964
; CURRENT FILING DATE: 2002-01-28
; EARLIER APPLICATION NUMBER: 09/314, 701
; EARLIER FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 241
; TYPE: PRT
; ORGANISM: p30-12
US-10-059-964-62

Alignment Scores:
Pred. No.: 48.8 Length: 241
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-20 (1-687) x US-10-059-964-62 (1-241)

OY 500 CTTAGTCTCCCAAGCAAAAC 520
|||||
Db 110 LeuSerSerProGlySerGluSer 116

RESULT 9
US-09-908-855-36
; Sequence 36, Application US/09908855
; Patent No. US20020115220A1
; GENERAL INFORMATION:
; APPLICANT: KONDO, KEIJI
; APPLICANT: MIURA, YUTAKA
; TITLE OF INVENTION: YEAST VECTOR AND METHOD OF PRODUCING PROTEINS USING THE
; FILE REFERENCE: 049441/0118
; CURRENT APPLICATION NUMBER: US/09/908, 855
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/242, 690
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Candida utilis
US-09-908-855-36

Alignment Scores:
Pred. No.: 47.9 Length: 267
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 10 Gaps: 0

US-09-835-992A-20 (1-687) x US-09-908-855-36 (1-267)

OY 507 GAACCTAGTACAGAAACACAT 487
|||||
Db 80 GluLeuSerLysLysHisAsn 86

RESULT 10
US-09-815-242-13249
; Sequence 13249, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815, 242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191, 078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206, 848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207, 727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242, 578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253, 625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257, 931

;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13249
;; LENGTH: 334
;; TYPE: PRF
;; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13249

Alignment Scores:
Pred. No.: 46.1 Length: 334
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-20 (1-687) x US-09-815-242-13249 (1-334)

QY 388 GCATTTCATCTATCAGAAC 368
|||||
Db 121 Alaphelbalaileargaaen 127

RESULT 11
US-09-205-658-110
;; Sequence 110, Application US/09205658
;; Patent No. US20010029617A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruvkun, Gary
;; APPLICANT: Ogg, Scott
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
;; FILE REFERENCE: 00786/351004
;; CURRENT APPLICATION NUMBER: US/09/205,658
;; CURRENT FILING DATE: 1998-12-03
;; EARLIER APPLICATION NUMBER: 08/857,076
;; EARLIER FILING DATE: 1997-05-15
;; EARLIER APPLICATION NUMBER: 08/888,534
;; EARLIER FILING DATE: 1997-07-07
;; EARLIER APPLICATION NUMBER: US98/10080
;; EARLIER FILING DATE: 1998-05-15
;; NUMBER OF SEQ ID NOS: 328
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 110
;; LENGTH: 380
;; TYPE: PRF
;; ORGANISM: Caenorhabditis elegans
US-09-205-658-110

Alignment Scores:
Pred. No.: 45 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-20 (1-687) x US-09-205-658-110 (1-380)

QY 94 CTAGTTTCTAATCTAATG 74
|||||
Db 4 Leuvalpheeluileumet 10

RESULT 12
US-09-844-353A-110
;; Sequence 110, Application US/09844353A
;; Patent No. US20020037585A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruvkun, Gary
;; APPLICANT: Kimura, Koufouou
;; APPLICANT: Patterson, Garth

;; APPLICANT: Ogg, Scott
;; APPLICANT: Paradis, Suzanne
;; APPLICANT: Tissenbaum, Heidi
;; APPLICANT: Morris, Jason
;; APPLICANT: Kowek, Allison
;; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
;; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
;; FILE REFERENCE: 00786/351005
;; CURRENT APPLICATION NUMBER: US/09/844,353A
;; CURRENT FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: US 08/857,076
;; PRIOR FILING DATE: 1997-05-15
;; NUMBER OF SEQ ID NOS: 114
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 110
;; LENGTH: 380
;; TYPE: PRF
;; ORGANISM: Caenorhabditis elegans
US-09-844-353A-110

Alignment Scores:
Pred. No.: 45 Length: 380
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-20 (1-687) x US-09-844-353A-110 (1-380)

QY 94 CTAGTTTCTAATCTAATG 74
|||||
Db 4 Leuvalpheeluileumet 10

RESULT 13
US-09-815-242-10888
;; Sequence 10888, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10888
;; LENGTH: 481
;; TYPE: PRF
;; ORGANISM: Enterococcus faecalis
US-09-815-242-10888

Alignment Scores:

Pred. No.:	43.2	Length:	481
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.30%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-815-242-10888 (1-481)

QY 191 AAACGTCATCGAAGCAA 211
|||||

Db 399 LysThAlaSerGluSerLys 405

RESULT 14
US-09-800-396-4

; Sequence 4, Application US/09800396
; Patent No. US2002005395A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola
; TITLE OF INVENTION: NOVEL RESPONSE REGULATOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Dechert Price & Rhoads
; STREET: 997 Lenox Drive, Building 3, Suite 210
; CITY: Lawrenceville
; STATE: NJ
; COUNTRY: USA
; ZIP: 08543
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/800,396
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/879,531
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bloom, Allen
; REGISTRATION NUMBER: 29,135
; REFERENCE/DOCKET NUMBER: GM10018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-520-3214
; TELEFAX: 609-520-3259
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-800-396-4

Alignment Scores:

Pred. No.:	42.1	Length:	554
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-800-396-4 (1-554)

QY 97 TTCCAGTTTTCTAATACTA 77
|||||

Db 6 PheLeuValPheLeuIleLeu 12

RESULT 15
US-09-925-299-859
; Sequence 859, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 859
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (590)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-859

Alignment Scores:

Pred. No.:	39.9	Length:	758
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	10	Gaps:	0

US-09-835-992A-20 (1-687) x US-09-925-299-859 (1-758)

QY 372 GAACAGTATTTCTTCCCAA 352
|||||

Db 151 GluThrValPheSerSerGln 157

Search completed: January 14, 2003, 18:08:18
Job time : 9.57721 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 ; Search time 23.0202 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992a-20

Perfect score: 212
Sequence: 1 atttaatttcataaac.....cttgtaagcacagttggt 687

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-O/cgpn2.1/USPTO.spool/US09835992/runat_14012003.161526.4038/arp_query.fasta.1.3932
-DB=pir_73 -QMT=fastan -SUFFIX=olig.rpr -MINMATCH=0.1 -DOPEL=0 -DOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09835992 @CGN.1.1.192 @runat_14012003.161526.4038 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MAIN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

PIR_73:*
1: PIR:*
2: PIR:*
3: PIR:*
4: PIR:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	268	2	F64024
2	8	3.8	342	2	A40191
3	8	3.8	343	2	T19082
4	8	3.8	393	2	H64233
5	8	3.8	512	2	E83060
6	8	3.8	542	2	S30884
7	8	3.8	791	2	H72552
8	8	3.8	1183	2	F90559
9	8	3.8	2329	2	T28125
10	7	3.3	55	2	G95972
11	7	3.3	65	2	S34538
12	7	3.3	91	2	T11269
13	7	3.3	105	2	E84075
14	7	3.3	120	2	T30769

15	7	3.3	126	2	F72150	BGL protein - vari
16	7	3.3	164	1	NW001	2S albumin 1 precu
17	7	3.3	164	1	NW003	2S albumin 3 precu
18	7	3.3	194	2	G69933	negative regulatio
19	7	3.3	195	2	E71000	hypothetical prote
20	7	3.3	201	2	G90674	probable transcript
21	7	3.3	201	2	B85525	probable transcript
22	7	3.3	201	2	F64758	yahd protein - Esc
23	7	3.3	205	2	T08398	calmodulin homolog
24	7	3.3	207	2	T24267	hypothetical prote
25	7	3.3	236	1	B64212	hypothetical prote
26	7	3.3	245	2	A30154	19e receptor alpha
27	7	3.3	254	2	D69140	hypothetical prote
28	7	3.3	269	2	AF1532	spermidine/putresc
29	7	3.3	269	2	AF1175	gamma-ABC transpor
30	7	3.3	271	2	F72379	sugar ABC transpor
31	7	3.3	276	2	AC1586	gamma-glutamyl kin
32	7	3.3	276	2	AD1232	gamma-glutamyl kin
33	7	3.3	278	2	C75616	oxidoreductase, sh
34	7	3.3	278	2	B75442	3-hydroxybutyryl-C
35	7	3.3	279	2	T20443	hypothetical prote
36	7	3.3	290	2	G72203	sugar ABC transpor
37	7	3.3	307	2	G64350	quinolinate synthe
38	7	3.3	313	2	D86829	ribose ABC transpo
39	7	3.3	314	2	T45010	probable site-spec
40	7	3.3	314	2	S73717	hypothetical prote
41	7	3.3	319	2	D90589	hypothetical prote
42	7	3.3	320	2	T30894	lipase homolog r0
43	7	3.3	326	1	VGXRKH	glycoprotein VP7 p
44	7	3.3	326	1	VGXRKH	glycoprotein VP7 p
45	7	3.3	326	1	A44891	glycoprotein VP7 p

ALIGNMENTS

RESULT 1

F64024 hypothetical protein H11273 - Haemophilus influenzae (strain Rd KW20)

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence=revision 18-Aug-1995 #text_change 08-Oct-1999

C:Accession: F64024

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kertlavage

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman

; D.M.; Brandon, R.C.; Fine, L.D.; Fritschman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vence

A:Title: Whole-genome random sequencing and assembly of Haemophilus Influenzae Rd.

A:Reference number: A64000; MUID:9350630; PMID:7542800

A:Accession: F64024

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-268 <TRGR>

A:Cross-references: GB:U32807; GB:L42023; NID:q1574725; PIDN:MAC22921.1; PID:q1574728

Alignment Scores:

Pred. No.: 10
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.79%
Indels: 0

US-09-835-992a-20 (1-687) x F64024 (1-268)

OY 522 GAGTTTCCTCGGAGAACTAAGT 499

DB 223 GlnPheSerLeuGlyGluLeuSer 230

RESULT 2

A40191

platelet-activating factor receptor - human

C:Species: Homo sapiens (man)

C:Date: 28-Aug-1992 #sequence=revision 28-Aug-1992 #text_change 20-Jun-2000

C:Accession: AA0191; JH0479; AA1079; JCI359; J48831; I51923
R:Kunz, D.; Gerard, N.P.; Gerard, C.
J. Biol. Chem. 267, 9101-9106, 1992
A:Title: The human leukocyte platelet-activating factor receptor. cDNA cloning, cell
A:Reference number: AA0191; MUID:92250505; PMID:1374385
A:Accession: AA0191
A:Molecule type: mRNA
A:Residues: 1-342 <KUN>
A:Cross-references: GB:M674; NID:g456293; PIDN:AAA60002.1; PID:g456294
R:Ye, R.D.; Prossnitz, E.R.; Zou, A.; Cochran, C.G.
Biochem. Biophys. Res. Commun. 180, 105-111, 1991
A:Title: Characterization of a human cDNA that encodes a functional receptor for platelet-
A:Reference number: JH0479; MUID:92028922; PMID:1656963
A:Accession: JH0479
A:Molecule type: mRNA
A:Residues: 1-342 <YER>
A:Cross-references: GB:M80436; NID:g189537; PIDN:AAA60001.1; PID:g189538
A:Experimental source: granulocyte, cell line HL-60 all
R:Nakamura, M.; Honda, Z.; Izumi, T.; Sakakura, C.; Mutoh, H.; Minami, M.; Bito, H.; Sey
J. Biol. Chem. 266, 20400-20405, 1991
A:Title: Molecular cloning and expression of platelet-activating factor receptor from hu
A:Reference number: AA1079; MUID:92041873; PMID:1657923
A:Accession: AA1079
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-342 <NAK>
A:Cross-references: GB:DI0202; GB:D90433; NID:g219975; PIDN:BA01050.1; PID:g219976
R:Suzumoto, T.; Tsuchimochi, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurachi, Y.
Biochem. Biophys. Res. Commun. 189, 617-624, 1992
A:Title: Molecular cloning and characterization of the platelet-activating factor recep
A:Reference number: JCI359; MUID:93112021; PMID:1281995
A:Accession: JCI359
A:Molecule type: mRNA
A:Residues: 1-315, 'N', 317-342 <SUG>
A:Experimental source: heart
A:Note: the authors translated the codon AAT for residue 316 as Lys
R:Seifried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W.
Genomics 13, 832-834, 1992
A:Title: The human platelet-activating factor receptor gene (PTAFR) contains no introns
A:Reference number: A4831; MUID:92347886; PMID:1322356
A:Accession: A4831
A:Molecule type: DNA
A:Residues: 1-226, 'TG', 229-342 <SEV>
A:Cross-references: GB:M8177; NID:g190697; PIDN:AAA60214.1; PID:g190698
A:Note: sequence extracted from NCBI backbone (NCBIN:109813, NCBI:109814)
R:Chase, P.B.; Halonene, M.; Regan, J.W.
Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
A:Title: Cloning of a human platelet-activating factor receptor gene: evidence for an in
A:Reference number: I51923; MUID:93192035; PMID:8383507
A:Accession: I51923
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <RES>
A:Cross-references: GB:S56396; NID:g298580; PIDN:AA025755.1; PID:g298581
A:Accession: I51923
A:Molecule type: DNA
A:Genetics:
A:Gene: GDB:PTAFR
A:Cross-references: GDB:128806; OMIM:173393
A:Map position: 1p35-1p34.3
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein
F:17-38/Domain: transmembrane #status predicted <TRI>
F:54-75/Domain: transmembrane #status predicted <TRI>
F:92-113/Domain: transmembrane #status predicted <TRI>
F:134-155/Domain: transmembrane #status predicted <TRI>
F:184-205/Domain: transmembrane #status predicted <TRI>
F:223-253/Domain: transmembrane #status predicted <TRI>
F:277-297/Domain: transmembrane #status predicted <TRI>
F:277-297/Domain: transmembrane #status predicted <TRI>
Alignment Scores:
Pred. No.: 9.81 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x A40191 (1-342)

OY 103 TCTTTTTCCTAGTTTCTAATA 80
    |||||||TCTTTTCTAATAA|||
DB 194 SerpPheLeuValPheLeuIle 201

RESULT 3
T19082
hypochemical protein C08B6.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19082
R:Wilkinson, J.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19070
A:Accession: T19082
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-343 <WIL>
A:Cross-references: EMBL:Z72502; PIDN:CAA96591.1; GSPDB:GN00023; CESP:C08B6.9
A:Experimental source: clone C08B6
C:Genetics:
A:Gene: CESP:C08B6.9
A:Map position: 5
A:Introns: 25/2; 84/3; 174/2; 290/2

Alignment Scores:
Pred. No.: 9.8 Length: 343
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x T19082 (1-343)

OY 628 CAGATAAAATTAGGACAGTAA 605
    |||||||TCTTTTCTAATAA|||
DB 23 GlnAsnLysIleArgAsnSerLys 30

RESULT 4
H64233
hypochemical protein MG306 - Mycoplasma genitalium
C:Species: Mycoplasma genitalium
C:Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 07-Dec-1999
C:Accession: H64233
R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, C.A.; Venter, J.C.
Science 270, 397-403, 1995
A:Title: The minimal gene complement of Mycoplasma genitalium.
A:Reference number: A64200; MUID:96026346; PMID:7563993
A:Accession: H64233
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <TIGR>
A:Cross-references: GB:U39712; GB:I43967; NID:q1046005; PID:q1046006; TIGR:MG306
A:Experimental source: strain G-37
C:Genetics:
A:Genetic code: SGC3

Alignment Scores:
Pred. No.: 9.68 Length: 393
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x H64233 (1-393)

```


OY 103 TCTTTTTCCTAGCTTTTCTAATA 80
 Db 19 SerpHeLeuValPheLeuIle 26

RESULT 5

E83060
 Iron (III)-transport system permease HltB PA4688 [imported] - Pseudomonas aeruginosa (str

C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: E83060

R:Stover, C.K.; Pham, X.Q.; Ertlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Llm,
 .; Loay, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: E83060

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 <STO>

A:Cross-references: GB:AE004882; GB:AE004091; NID:9950939; PIDN:AG08075.1; GSPDB:G0001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: hltB; PA4688

C:Superfamily: stub protein

Alignment Scores:

Pred. No.:	9.45	Length:	512
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-20 (1-687) x E83060 (1-512)

OY 479 CTCTTGGCAATGCTTTCTTACTT 502

Db 230 LeuValLeuGluValPheLeuIle 237

RESULT 6

S30884
 SRP1 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: Protein N1606; Protein YNL189W

C:Species: Saccharomyces cerevisiae

C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Jun-2000

C:Accession: S30884; S63144

R:Tano, R.; Oakes, M.; Yamagishi, M.; Dodd, J.A.; Nomura, M.

Mol. Cell. Biol. 12, 5640-5651, 1992

A:Title: Cloning and characterization of SRP1, a suppressor of temperature-sensitive RN

A:Reference number: S30884; MUID:93078765; PMID:1448093

A:Accession: S30884

A:Molecule type: DNA

A:Residues: 1-542 <YAN>

A:Cross-references: EMBL:M5849; NID:9172702; PIDN:AAA35090.1; PID:9172703

R:Oermier, B.; Pirevandi, E.; Rinke, M.; Domdey, H.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63122

A:Accession: S63144

A:Molecule type: DNA

A:Residues: 1-542 <OBE>

A:Cross-references: EMBL:Z71465; NID:91302178; PIDN:CAA96083.1; PID:91302179; MIPS:YNL18

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SCD:SRP1; SCM1; KAP1

A:Cross-references: SGD:S0005133; MIPS:YNL189W

A:Map position: 14L

C:Superfamily: pendulin

C:Keywords: nucleus; transmembrane protein

F:336-352/Domain: transmembrane #status predicted <TM>

Alignment Scores:

Pred. No.:	9.4	Length:	542
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-20 (1-687) x S30884 (1-542)

OY 497 TTACTAGTCTCCAGGAAAC 520

Db 349 LeuSerSerProLysGluAsn 356

RESULT 7

H72552
 hypothetical protein APE1708 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: H72552

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310335; PMID:10382366

A:Accession: H72552

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-791 <RAM>

A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BA80709.1; PID:01044495; PID:9

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1708

Alignment Scores:

Pred. No.:	9.07	Length:	791
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.77%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-20 (1-687) x H72552 (1-791)

OY 76 TTAGTATTGAAAACTAGGAAA 99

Db 134 LeuValLeuGluLysLeuGlyLys 141

RESULT 8

F90559
 conserved hypothetical protein MYPu_3820 [imported] - Mycoplasma pulmonis (strain UAB

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001

C:Accession: F90559

R:Chambud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer,

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: F90559

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1183 <RUR>

A:Cross-references: GB:AL445566; PID:914089796; PIDN:CAC13555.1; GSPDB:G000153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYPu_3820

A:Genetic code: SGC3

Alignment Scores:

Pred. No.:	8.74	Length:	1183
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.79%	Indels:	0

DB: 2 Gaps: 0
US-09-835-992a-20 (1-687) x F90559 (1-1183)
OY 194 TTTTCAATATATATCATTTCA 171
|||||
DB 577 PheSerAsnTyrAsnLeuIleSer 584
RESULT 9
T28125
hypothetical protein ZK945.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T28125
R:Wilkinson, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z20472
A:Accession: T28125
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2329 <MIL>
A:Cross-references: EMBL:Z48544; NID:g9695493; PIDN:CAA88442.1; GSPDB:GN00020; CESP:ZK945
A:Experimental source: clone ZK945
C:Genetics:
A:Gene: CESP:ZK945.9
A:Map position: 2
A:Introns: 160/3; 187/1; 243/1; 580/3; 607/3; 705/3; 789/1; 864/3; 940/3; 974/3; 1064/2;
Alignment Scores:
Pred. No.: 8 21 Length: 2329
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 2 Gaps: 0
US-09-835-992a-20 (1-687) x T28125 (1-2329)
OY 398 AGATTTTATCTCNCCTAAATTTTA 421
|||||
DB 1632 ArgPheLeuIleSerLeuIleLeu 1639
RESULT 10
G95972
hypothetical protein [imported] - *Sinorhizobium meliloti* (strain 1021) megaplasmid pSymH
C:Species: *Sinorhizobium meliloti*
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: G95972
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: G95972
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-55 <KUR>
A:Cross-references: GB:AL591985; PIDN:GAC49447.1; PID:g15140933; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
peta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The complete genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Accession: T11269
A:Contents: annotation
C:Genetics:
A:Gene: SMD1672
A:Genome: plasmid
Alignment Scores:
Pred. No.: 136 Length: 55

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 2 Gaps: 0
US-09-835-992a-20 (1-687) x G95972 (1-55)
OY 529 AGTTTAAGAGTTCCTTGGG 509
|||||
DB 25 SerLeuArgValPheLeuIle 51
RESULT 11
S34538
hypothetical protein 65 (rps18 3' region) - *Euglena gracilis* chloroplast
C:Species: chloroplast *Euglena gracilis*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Sep-1999
C:Accession: S34538, S34905
R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.; Monfort, A.; Orsat, B.; Spielma
submitted to the EMBL Data Library, January 1993
A:Description: The complete sequence of the *Euglena gracilis* chloroplast genome (tent
A:Reference number: S34494
A:Accession: S34538
A:Molecule type: DNA
A:Residues: 1-65 <HAL1>
A:Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774
R:Hallick, R.B.; Hong, L.; Drager, R.G.; Favreau, M.R.; Monfort, A.; Orsat, B.; Spiel
Nucleic Acids Res. 21, 3537-3544, 1993
A:Title: Complete sequence of *Euglena gracilis* chloroplast DNA.
A:Reference number: S34862; MUID:93347989; PMID:8346031
A:Accession: S34905
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-65 <HAL2>
A:Cross-references: EMBL:X70810; NID:g415327; PIDN:CAA50118.1; PID:g415774
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Genetics:
A:Superfamily: conserved hypothetical protein ycf9
C:Keywords: chloroplast
Alignment Scores:
Pred. No.: 134 Length: 65
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 2 Gaps: 0
US-09-835-992a-20 (1-687) x S34538 (1-65)
OY 100 TTTTCTAGTTTCTATA 80
|||||
DB 51 PhePheLeuValPheLeuIle 57
RESULT 12
T11269
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4L - hardbacked tick (*Ixodes hexa*
C:Species: mitochondrion *Ixodes hexagonus*
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 03-Jun-2002
C:Accession: T11269
R:Black IV, W.C.; Roehrdanz, R.L.
Mol. Biol. Evol. 15, 1772-1785, 1998
A:Title: Mitochondrial gene order is not conserved in arthropods: prostrate and meta
A:Reference number: Z17252; MUID:99083443; PMID:9866211
A:Accession: T11269
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-91 <BLA>
A:Cross-references: EMBL:AF081828; NID:g4164542; PID:g4164552; PIDN:AAD05513.1
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4L

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
Pred. No.: 130 Length: 91
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x T11269 (1-91)

OY 225 TTATTCCTTAATGATCGCTTT 245

DB 37 LeuLeuPheLeuLeuLeuPhe 43

RESULT 13

E84075
hypothetical protein BH3405 [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E84075
R:Takami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E84075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <STO>
A:Cross-references: GB:AP001518; GB:BA000004; MID:g10175792; PIDN:BA07124.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3405

Alignment Scores:
Pred. No.: 128 Length: 105
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x E84075 (1-105)

OY 606 TTACTGTTCTAATTATTC 626

DB 11 LeuLeuPheLeuLeuPhe 17

RESULT 14

T30769
hypothetical protein 8L - Vaccinia virus (strain Ankara)
N:Alternate names: putative 13.7k protein
C:Species: vaccinia virus
A:Variety: strain Ankara
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R:Antoine, G.; Schefflinger, F.; Falkner, F.G.; Dorner, F.
C:Accession: T30769
submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A:Reference number: Z20877
A:Accession: T30769
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-120 <ANT>
A:Cross-references: EMBL:U94848; PIDN:AAB96403.1
A:Experimental source: strain Ankara
C:Genetics:
A:Note: MVA008L

Alignment Scores:
Pred. No.: 126 Length: 120
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x T30769 (1-120)

OY 360 AAAATACGTTCGATAGCA 380

DB 2 LysileuPheLeuLeuLea 8

RESULT 15

F72150
B6L protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 24-Nov-1999
C:Accession: F72150
R:Shchekunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safonov, P.F.; Masung, R.F.; Lo
submitted to Genbank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m
A:Reference number: A72150
A:Accession: F72150
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <SHC>
A:Cross-references: GB:Y16780; MID:g5830555; PIDN:CAB54600.1; PID:el542556; PID:g5830
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: B6L

Alignment Scores:
Pred. No.: 126 Length: 126
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-20 (1-687) x F72150 (1-126)

OY 360 AAAATACGTTCGATAGCA 380

DB 2 LysileuPheLeuLeuLea 8

Search completed: January 14, 2003, 17:40:39
Job time: 27.0202 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 10.6084 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992A-20

Perfect score: 212
Sequence: 1 atttaatttctatataaac.....cttgttalcgcacagttgt 687

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p.model -DEV=xlp
-O=/gen2.1/USPTO.spool/US09835992/runat.14012003.161525.4014/app.query.fasta.1.3932
-DB=SwissProt_40 -QMT=faasan -SUFFIX=olig.rsp -MINMATCH=0.1 -DOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.coi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFM=pro -NORW-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992.@CGN.1.1.77.@runat.14012003.161525.4014 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NRC_SCORES=0 -WAIT -LONGLOC -DEV_TIMELIMIT=120
-MARN_TIMELIMIT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FAPOP=6 -FAPEXT=7
-YAPOP=60 -YAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	3.8	208	1	PAFR_MACMU
2	8	3.8	268	1	YC73_HAEIN
3	8	3.8	342	1	PAFR_HUMAN
4	8	3.8	393	1	Y306_MYCE
5	8	3.8	542	1	Y306_MYCE
6	8	3.8	3178	1	Y306_MYCE
7	8	3.3	65	1	Y306_MYCE
8	7	3.3	164	1	Y306_MYCE
9	7	3.3	164	1	Y306_MYCE
10	7	3.3	180	1	Y306_MYCE
11	7	3.3	194	1	Y306_MYCE
12	7	3.3	201	1	Y306_MYCE
13	7	3.3	245	1	Y306_MYCE
14	7	3.3	267	1	Y306_MYCE
15	7	3.3	268	1	Y306_MYCE
16	7	3.3	276	1	Y306_MYCE
17	7	3.3	276	1	Y306_MYCE
18	7	3.3	278	1	Y306_MYCE

C	19	7	3.3	278	1	Y110_MYCE	P47356 mycoplasma
	20	7	3.3	307	1	NADA_METJA	O57850 methanococ
	21	7	3.3	314	1	Y315_MYCPA	O50362 mycoplasma
	22	7	3.3	326	1	VS09_ROT2	O03874 equine rota
	23	7	3.3	326	1	VS09_ROT2	P18855 human rota
	24	7	3.3	326	1	VS09_ROT2	P12476 thesus rota
	25	7	3.3	345	1	CLT2_PIG	O59033 sus scrofa
	26	7	3.3	372	1	FLGI_BUCAI	P57426 buchnera ap
	27	7	3.3	379	1	PANE_YEAST	P38787 saccharomyc
	28	7	3.3	380	1	CAPM_STPAU	P39862 staphylococ
	29	7	3.3	405	1	DXR_PASMU	P57995 pasteurella
	30	7	3.3	423	1	MTB5_NERGO	O59605 neisseria 1
	31	7	3.3	423	1	MTN4_NEILA	P50182 neisseria 1
	32	7	3.3	473	1	Y099_CAMEL	P37821 malus domes
	33	7	3.3	527	1	Y099_CAMEL	P41849 caenorhabd1
	34	7	3.3	536	1	IMAS_HUMAN	O51511 homo sapien
	35	7	3.3	538	1	IMAI_HUMAN	P52284 homo sapien
	36	7	3.3	538	1	IMAI_HUMAN	O09660 mus musculu
	37	7	3.3	551	1	VPI7_YEAST	P32913 saccharomyc
	38	7	3.3	560	1	YJK5_YEAST	P42948 saccharomyc
	39	7	3.3	584	1	LEU1_SCHPO	O59736 schizosacch
	40	7	3.3	613	1	UL06_EBV	P03213 Epstein-Bar
	41	7	3.3	681	1	VGP_MABVM	P35253 marburg vir
	42	7	3.3	688	1	YAOX_SCHPO	O10109 schizosacch
	43	7	3.3	690	1	SYGB_BUCAI	P57235 buchnera ap
	44	7	3.3	778	1	RHP9_SCHPO	P87074 schizosacch
	45	7	3.3	1068	1	YCFO_MARPO	P12221 marichantia

ALIGNMENTS

RESULT 1
PAFR_MACMU STANDARD; PRT: 208 AA.
ID PAFR_MACMU
AC P35366:
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
ID Platelet activating factor receptor (PAF-R) (Fragment).
GN PAFR.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RL Behal R.H., Debussche M.S., Olson M.S.;
RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: RECEPTOR FOR PLATELET ACTIVATING FACTOR, A CHEMOTACTIC
CC PHOSPHOLIPID MEDIATOR THAT POSSESSES POTENT INFLAMMATORY, SMOOTH-
CC MUSCLE CONTRACTIONS AND HYPERTENSIVE ACTIVITY. SEEM TO MEDIATE ITS
CC ACTION VIA A G PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L07333; AAA6893.1;
CC InterPro: IPR000276; GPCR_Rhodopsin.
CC Pfam: PF00001; 7tm_1; 1.
CC PROSITE: PS00237; G_PROTEIN_RECPT_P1_1; 1.
CC PROSITE: PS00262; G_PROTEIN_RECPT_P1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Chemotaxis.
CC DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
CC FT

```

FT TRANSMEM 17 38 1 (POTENTIAL).
FT DOMAIN 39 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 74 2 (POTENTIAL).
FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 92 113 3 (POTENTIAL).
FT DOMAIN 114 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 155 4 (POTENTIAL).
FT DOMAIN 156 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 >208 CYTOPLASMIC (POTENTIAL).
FT CAROXYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 90 173 BY SIMILARITY.
FT NON_TER 208
SQ SEQUENCE 208 AA; 23736 MW; 7985928F70B3C6A1 CRC64;

```

Alignment Scores:

```

Pred. No.: 9.48 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0

```

US-09-835-992a-20 (1-687) x PAFR_MACMU (1-208)

```

QY 103 TCTTTTCCAGTTTCTTAATA 80
DB 194 SerpPhelLeuValPheLeuLe 201

```

RESULT 2

```

YC73_HAEIN STANDARD: PRT; 268 AA.
ID YC73_HAEIN
AC P44150;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11273.
GN H11273.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Furmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).

```

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U32807; AAC22921.1; -
 CC DR TIGR; H11273; -
 CC DR InterPro; IPR000051; SAM_bind.
 CC KW Hypothetical protein; Complete proteome.
 CC SEQUENCE 268 AA; 30510 MW; E5B2BDA7AAD4D0B CRC64;

```

Alignment Scores:
Pred. No.: 9.03 Length: 268
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.79% Indels: 0
DB: 1 Gaps: 0

```

US-09-835-992a-20 (1-687) x YC73_HAEIN (1-268)

```

QY 522 GAGTTTCCTGGAGAGACTAGT 499
DB 223 GlupheserLeuGlyLeuSer 230

```

RESULT 3

```

PAFR_HUMAN STANDARD: PRT; 342 AA.
ID PAFR_HUMAN
AC P25105;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Platelet activating factor receptor (PAF-R).
GN PAFR OR PAFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte; PubMed=1656963;
RX MEDLINE=92028922; PubMed=1656963;
RA Ye R.D., Prossnitz E.R., Zou A., Cochrane C.G.;
RT "Characterization of a human cDNA that encodes a functional receptor
RT for platelet activating factor."
RL Biochem. Biophys. Res. Commun. 180:105-111(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte; PubMed=1657923;
RX MEDLINE=92041873; PubMed=1657923;
RA Nakamura M., Honda Z., Izumi T., Sakanaka C., Mutoh H., Minami M.,
RA Bito H., Seyama Y., Matsumoto T., Noma M., Shimizu T.;
RT "Molecular cloning and expression of platelet-activating factor
RT receptor from human leukocytes."
RL J. Biol. Chem. 266:20400-20405(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92250505; PubMed=1374385;
RA Kunz D., Gerard N.P., Gerard C.;
RT "The human leukocyte platelet-activating factor receptor. cDNA
RT cloning, cell surface expression, and construction of a novel
RT epitope-bearing analog."
RL J. Biol. Chem. 267:9101-9106(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347886; PubMed=1322356;
RA Seyfried C.E., Schweikart V.L., Godiska R., Gray P.W.;
RT "The human platelet-activating factor receptor gene (PTAFR) contains
RT no introns and maps to chromosome 1."
RL Genomics 13:832-834(1992).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart; PubMed=1281995;
RX MEDLINE=93112021; PubMed=1281995;
RA Sugimoto T., Tsuchimochi H., McGregor C.G., Mutoh H., Shimizu T.,
RA Kurachi Y.;
RT "Molecular cloning and characterization of the platelet-activating
RT factor receptor gene expressed in the human heart."
RL Biochem. Biophys. Res. Commun. 189:617-624(1992).
RN [6]
RP SEQUENCE FROM N.A.
RA Behal R.H., Debussere M.S., Olson M.S.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.

```


ID IMAL YEAST STANDARD: PRT: 542 AA.
 AC 002821;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Importin alpha subunit (Karyopherin alpha subunit) (Serine-rich RNA
 DE polymerase I suppressor protein).
 GN SRP1 OR KAP60 OR YNL19W OR N1606.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 NC NCB1_TaxID=4932;
 RX MEDLINE-93078765; PubMed-1448093;
 RA Yano R., Oakes M., Yamagishi M., Dodd J.A., Nomura M.;
 RT "Cloning and characterization of SRP1, a suppressor of temperature-
 RT sensitive RNA polymerase I mutations, in Saccharomyces cerevisiae.";
 RL Mol. Cell. Biol. 12:5640-5651(1992).
 RN [2]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RC STRAIN-W303;
 RX MEDLINE-96004702; PubMed-7565597;
 RA Kuesel P., Frasch M.;
 RT "Yeast Srp1, a nuclear protein related to Drosophila and mouse
 RT penulin, is required for normal mitrigration, division, and integrity
 RT of nuclei during mitosis.";
 RL Mol. Gen. Genet. 248:351-363(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
 RL Submitted (May-1996) to the EMBL/GenBank/DBD databases.
 RN [4]
 RP MUTAGENESIS.
 RX MEDLINE-94316601; PubMed-8041713;
 RA Yano R., Oakes M.L., Tabb M.M., Nomura M.;
 RT "Yeast Srp1p has homology to armadillo/plakoglobin/beta-catenin and
 RT participates in apparently multiple nuclear functions including the
 RT maintenance of the nucleolar structure.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6880-6884(1994).
 RN [5]
 RP IDENTIFICATION.
 RX MEDLINE-95348058; PubMed-7622450;
 RA Enekel C., Blobel G., Rexach M.;
 RT "Identification of a yeast karyopherin heterodimer that targets
 RT import substrate to mammalian nuclear pore complexes.";
 RL J. Biol. Chem. 270:16499-16502(1995).
 RN [6]
 RP NUCLEOPORIN REPEAT BINDING REQUIREMENT.
 RX MEDLINE-96097112; PubMed-8521485;
 RA Rexach M., Blobel G.;
 RT "Protein import into nuclei: association and dissociation reactions
 RT involving transport substrate, transport factors, and nucleoporins.";
 RL Cell 83:683-692(1995).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 89-510.
 RX MEDLINE-98359119; PubMed-9695948;
 RA Conti E., Uy M., Leighton L., Blobel G., Kuriyan J.;
 RT "Crystallographic analysis of the recognition of a nuclear
 RT localization signal by the nuclear import factor karyopherin alpha.";
 RL Cell 94:193-204(1998).
 CC -1- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
 CC EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
 CC SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
 CC RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS A COMPLEX WITH AN IMPORTIN BETA SUBUNIT. INTERACTS
 CC WITH CSE1.
 CC -1- SUBCELLULAR LOCATION: MAINLY LOCALIZED AT THE PERIPHERY OF THE
 CC NUCLEUS.
 CC -1- MISCELLANEOUS: BINDS TO NUCLEOPORIN FXFG BUT NOT GLFG REPEAT
 CC REGIONS. RAN-GTP CAN DISRUPT THE KARYOPHERIN HETERODIMER BY
 CC BINDING TO THE BETA SUBUNIT AND RELEASES BOTH SUBUNITS FROM THE
 CC DOCKING SITE.

CC -1- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: M75849; AAA5090.1; -;
 DR EMBL: Z71465; CA936083.1; -;
 DR PIR: S30884; S30884.
 DR PDB: 1BK5; 16-FEB-99.
 DR PDB: 1BK6; 13-JAN-99.
 DR SGD: S0005133; SRP1.
 DR InterPro: IPR000225; Armadillo.
 DR InterPro: IPR002652; Importina_B.
 DR Pfam: PF00514; Armadillo_seg; 8.
 DR Pfam: PF01749; IBB; 1.
 DR SMART: SM00185; ARM; 8.
 DR PROSITE: PS50176; ARM_REPEAT; 2.
 DR Transprot; Protein transport; Repeat; Nuclear protein; 3D-structure.
 FT DOMAIN 17 38
 FT REPEAT 123 165 IBB.
 FT REPEAT 166 207 ARM 1.
 FT REPEAT 208 249 ARM 2.
 FT REPEAT 250 291 ARM 3.
 FT REPEAT 292 333 ARM 4.
 FT REPEAT 334 375 ARM 5.
 FT REPEAT 376 417 ARM 6.
 FT REPEAT 418 462 ARM 7.
 FT DOMAIN 463 542 ARM 8.
 FT MUTAGEN 116 116 ASP/GLU-RICH (ACIDIC).
 FT S->F: IN SRP1-31; TS MUTANT; REDUCED
 FT GROWTH RATE AND CHROMOSOME LOSS.
 FT E->K: IN SRP1-49; TS MUTANT; ALTERATION
 FT IN NUCLEOLAR AND MICROTUBULE MORPHOLOGY.
 FT P->Q: IN SRP1-1; TS MUTANT.
 FT D->N: IN SRP1-3; TS MUTANT.
 FT E->K: IN SRP1-2; TS MUTANT.
 FT G->V: IN SRP1-54; TS MUTANT; REDUCED
 FT GROWTH RATE.
 SQ SEQUENCE 542 AA; 60441 MW; 8D3A0CB76F2E7C00 CRC64;
 Alignment Scores:
 Pred. No.: 7 88 Length: 542
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.77% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992A-20 (1-687) x IMAL_YEAST (1-542)
 QY 497 TTAAGTTAGTTCGCCAAGGAAGAAC 520
 DB 349 LeuLeuSerSerProLySGluLsn 356
 ||||||||||||||||||||
 RESULT 6
 YS89_CABEEL
 ID YS89_CABEEL STANDARD: PRT: 3178 AA.
 AC 009624; 009625; 0969D4;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK945.9 in chromosome II.
 GN ZK945.9/ZK945.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NC NCB1_TaxID=6239;
 RN [1]


```

RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Wilkison-Sprot J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RL Durbin R.;
RC Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48544; CAB70192.1; -.
DR EMBL: Z48582; CAB70192.1; JOINED.
DR EMBL: Z48544; CAB70201.1; JOINED.
DR WormPep: ZK945.9; CE25697.
DR InterPro: IPR002111; Cat_channel_TrpL.
DR InterPro: IPR001024; Lipoxigenase_LH2.
DR InterPro: IPR000636; M+channel_nlg.
DR InterPro: IPR000203; PKD_cys_rich.
DR Pfam: PF00520; Ion_trans; 1.
DR Pfam: PF01477; PLAT; 1.
DR Pfam: PF01825; GPS; 1.
DR SMART: SM00303; GPS; 1.
DR SMART: SM00308; LH2; 1.
KM Hypothetical protein: Transmembrane.
FT DOMAIN 265 1196 SER/TRP-RICH.
FT DOMAIN 1105 1241 GLY/SER-RICH.
FT DOMAIN 2071 2120 GPS.
FT TRANSMEM 13 30 POTENTIAL.
FT TRANSMEM 51 73 POTENTIAL.
FT TRANSMEM 2139 2161 POTENTIAL.
FT TRANSMEM 2348 2367 POTENTIAL.
FT TRANSMEM 2390 2412 POTENTIAL.
FT TRANSMEM 2451 2468 POTENTIAL.
FT TRANSMEM 2483 2505 POTENTIAL.
FT TRANSMEM 2567 2589 POTENTIAL.
FT TRANSMEM 2836 2858 POTENTIAL.
FT TRANSMEM 2939 2961 POTENTIAL.
FT TRANSMEM 2976 2996 POTENTIAL.
FT TRANSMEM 3038 3060 POTENTIAL.
SQ SEQUENCE 3178 AA; 344726 MW; F8239436D03666CD CRC64;

Alignment Scores:
Pred. No.: 5 61 Length: 3178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.77% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-20 (1-687) x YS89_CAEEL (1-3178)
QY 398 AGATTTTAACTCNCATATTTTA 421
DB 2481 ArgPheLeuIleSerLeuIleLeu 2488

RESULT 7
YCF9_EUGGR STANDARD; PRT; 65 AA.
AC P32095;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 7.6 kDa protein ycf9 (ORF 65).

```

```

GN YCF9.
OS Euglena gracilis.
OG Chloroplast.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Z;
RX MDLINE=93347989; PubMed=8346031;
RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
RA Orsat B., Spielmann A., Stutz E.;
RL "Complete sequence of Euglena gracilis chloroplast DNA.";
CC Nucleic Acids Res. 21:3537-3544(1993).
CC -1- SIMILARITY: BELONGS TO THE YCF9 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z11874; -. NOT_ANNOTATED_CDS.
DR EMBL: X70810; CAA50118.1; -.
DR PIR: S34538; S34538.
DR InterPro: IPR002644; ycf9_struct.
DR Pfam: PF01737; YCF9; 1.
DR ProDom: PD004770; ycf9_struct; 1.
KM Chloroplast: Hypothetical protein.
SQ SEQUENCE 65 AA; 7586 MW; 67DE2359D7A0F771 CRC64;

Alignment Scores:
Pred. No.: 120 Length: 65
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-20 (1-687) x YCF9_EUGGR (1-65)
QY 100 TTTTCTAGTTTCTGATTA 80
DB 51 PhePheLeuValPheLeuIle 57

RESULT 8
ZSS1_ARATH STANDARD; PRT; 164 AA.
ID ZSS1_ARATH
AC P15457;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2S seed storage protein 1 precursor (2S albumin storage protein)
DE (NM002-2S albumin 1).
GN AT2S1 OR AT4G27140 OR T24A18.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-73 AND 84-162.
RC STRAIN-cv. C24;
RA Krebbers E., Herdies L., de Clercq A., Seurinck J., Leemans J.,
RA van Damme J., Segura M., Gheysen G., van Montagu M.,
RA Vandekerckhove J.;
RT "Determination of the processing sites of an Arabidopsis 2S albumin
RL and characterization of the complete gene family.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-cv. C24;

```

RA Conceicao A.D.S., Krebbers E.;
 RL Submitted (JUL-1993) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Columbia;
 RX MEDLINE-20083488; PubMed-10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaler B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Pulgomech P., Watson M., Schmidheini T.,
 RA Reihert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohelsel J., Zimmermann W., Medler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Billam L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weltenegeger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., Van Staveren M., Dirse P.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysmaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Pettitt A., Rajandream M.A., Lyne M., Benes V., Reumann S.,
 RA Borkova D., Bioecker H., Scharfe M., Grimm M., Loehnett T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argilou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parrell L., Dedha N., Gnoj L., Schurz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Laitelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Min P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Yil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Rodriguez M., Hoffman J., Till S.,
 RA Chen E., Mariz M., Martensen R., McCombie W.R.,
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: THIS IS THE MOST ABUNDANT ISOFORM OF 2S ALBUMIN IN
 CC ARABIDOPSIS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; M22032; AAA32743.1;
 DR EMBL; Z24745; CAA80870.1;
 DR EMBL; A13820; CAA01132.1;
 DR EMBL; AL035680; CAB8844.1;
 DR EMBL; AL161565; CAB79569.1;
 DR PIR; JA0161; NMM01.

DR PIR; S34676; S34676.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000617; Napin.
 DR InterPro; IPR001768; TRY/amy1_inhbr.
 DR Pfam; PF00234; tryp_alpha_aml1; 1.
 DR PRINTS; PD00496; NAPIN.
 DR PRODOM; PD002498; Napin; 1.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Albumin; signal; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 37
 FT CHAIN 38 73
 FT PROPEP 74 83
 FT CHAIN 84 162
 FT PROPEP 163 164
 SQ SEQUENCE 164 AA; 19014 MW; 2BF28CB47AD9832B CRC64;
 Alignment Scores:
 Pred. No.: 101
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.30%
 DB: 1
 Gaps: 0
 Length: 164
 Matches: 7
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0
 US-09-835-992a-20 (1-687) x 2SSI_ARATH (1-164)
 QY 482 TTGGCATGTGTTCTTACTT 502
 DB 12 LeuAlaLeucyspheuleu 18
 RESULT 9
 ID 2SS3_ARATH STANDARD; PRT; 164 AA.
 AC P15459;
 DT 01-APR-1990 (rel. 14, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 25 seed storage protein 3 precursor (2S albumin storage protein)
 DE (NMW02-25 albumin 3).
 GN AT253 OR AT4G27160 OR T24A18.110.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID-3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA van Damme J., Segura M., Gheysen G., van Montagu M.,
 RA Vandekerckhove J.;
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Conceicao A.D.S., Krebbers E.;
 RL Submitted (JUL-1993) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. Columbia;
 RX MEDLINE-20083488; PubMed-10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaler B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Pulgomech P., Watson M., Schmidheini T.,
 RA Reihert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohelsel J., Zimmermann W., Medler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Billam L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,

ID YPBH_BACSU STANDARD; PRT; 194 AA.
AC P50734;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ypbh.
GN YPBH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RA MEDLINE=96349105; PubMed=8760912;
RX Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RT Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
the *serA* and *kds* loci cloned in a yeast artificial chromosome.";
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RX Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolojin A., Borchert S.,
RA Borriss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Goldschly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kraerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche R., Rose M., Sadale I.,
RA Sato T., Scanlan E., Schleich S., Schroeder R., Scrofano F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosto V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Vari A., Wambut R., Medler E., Medler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.";
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; L47648; AAC83952.1; -
DR EMBL; 299115; CAB14213.1; -
DR EMBL; 299116; CAB14229.1; -
DR Subtilisin; BG11434; ypbh.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 194 AA; 22157 MW; B27C6ABEA0745C CRC64;

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.32%
DB: 1
Gaps: 0
US-09-835-992a-20 (1-687) x YPBH_BACSU (1-194)
QY 303 AAAATAAAATATTTTAAACA 283
Db 10 Lysiletyllepheleuthr 16
RESULT 12
YAHB_ECOLI
ID YAHB_ECOLI STANDARD; PRT; 201 AA.
AC P77736;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical ANK-repeats protein yahb.
GN YAHB OR B0318.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Duncan M., Allen E., Araujo R., Aparicio A.M., Chung E., Davis K.,
RA Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O., Lew H., Lin D.,
RA Namath A., Oefner P., Roberts D., Schramm S., Davis R.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL; AE000139; AAC73421.1; -
DR EMBL; U73857; AAB18044.1; -
DR Ecocore; EGI3588; YAHB.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 6.
DR SMART; SM00248; ANK; 2.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW Hypothetical protein; Repeat; ANK repeat; Complete proteome.
FT REPEAT 5 34
FT REPEAT 38 67
FT REPEAT 71 100
FT REPEAT 104 134
FT REPEAT 138 172
FT REPEAT 176 201
SQ SEQUENCE 201 AA; 21687 MW; A90244CEB892E8BD CRC64;

Alignment Scores:
Pred. No.: 96.7
Score: 7.00
Length: 201
Matches: 7
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.30%
Indels: 0

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U08629; AAA65978.1; -.
 DR HSSP: P03962; IDQW.
 DR InterPro: IPR001754; OMPdecase.
 DR Pfam: PF00215; OMPdecase; 1.
 DR PROSITE: PS00156; OMPDECASE; 1.
 KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
 FT ACT_SITE 94 BY SIMILARITY.
 SQ SEQUENCE 267 AA; 29529 MW; D98361C398F49E06 CRC64;

Alignment Scores:

Pred. No.:	91.6	Length:	267
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992a-20 (1-687) x PYRF_PICST (1-267)

QY 507 GAACCTAGTAAGAACACCAAT 487

Db 80 GluLeuSerIysIysHisAsn 86

RESULT 15

PYRF_CANTR STANDARD; PRT; 268 AA.

AC 042771;
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)
 DE (OMPCase) (OMPdecase) (Uridine 5'-monophosphate synthase) (UMP
 DE synthase).
 GN URA3
 OS Candida tropicalis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M4;
 RX MEDLINE=98403409; PubMed=9734982;
 RA Su J.-H., Hsia J.-H., Chang M.-C.;
 RT "Cloning and sequence analysis of the Candida tropicalis URA3 gene
 RT encoding orotidine-5'-phosphate decarboxylase.";
 RL Curr. Microbiol. 37:210-213(1998)
 CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).
 CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
 CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF040702; AAB96773.1; -.
 DR HSSP: P03962; IDQW.
 DR InterPro: IPR001754; OMPdecase.
 DR Pfam: PF00215; OMPdecase; 1.
 DR PROSITE: PS00156; OMPDECASE; 1.
 KW Pyrimidine biosynthesis; Lyase; Decarboxylase.
 FT ACT_SITE 94 BY SIMILARITY.
 SQ SEQUENCE 268 AA; 29672 MW; A43CCIEDICH24237 CRC64;

Alignment Scores:

Pred. No.:	91.5	Length:	268
Score:	7.00	Matches:	7

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	1	Gaps:	0

US-09-835-992a-20 (1-687) x PYRF_CANTR (1-268)

QY 507 GAACCTAGTAAGAACACCAAT 487

Db 80 GluLeuSerIysIysHisAsn 86

Search completed: January 14, 2003, 17:29:11
 Job time: 14.6084 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 48.3743 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992a-20

Perfect score: 212
Sequence: 1 attttaatttcatttaaac.....cttgatgacacagttgt 687

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Labeling first 45 summaries

Command line parameters:

-MODEL-frame+ n2p.model -DEV-xip
-O/cgpn2.1/USPTO-US09835992/runat_14012003_161526_4027/app_query.fasta_1.3932
-DB-SPREMBL.21 -QFMT-fastan -SUFFIX-olg.rspt -MIMATCH=0.1 -DOOPCL=0
-LOOEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-olgigo -TRANS-human40.cdi
-LIST=45 -DOCALLIN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-OUTFMT=pic -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09835992 -CGN.1.1.346 -runat_14012003_161526_4027 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FAPOP=6 -FAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database:

1: SP_Archea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_Virus:*
16: SP_Bacteriophage:*
17: SP_Archaeo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	8	3.8	151 6 097662	097662 ovls arles

2	8	3.8	280 5	09N5G4	09N5G4 caenorhabd
C 3	8	3.8	342 6	09GK76	09GK76 capra hircu
C 4	8	3.8	342 6	09TRV5	09TRV5 bos taurus
C 5	8	3.8	343 5	017820	017820 caenorhabd
C 6	8	3.8	352 17	08TRV25	08TRV25 methanosarc
C 7	8	3.8	376 16	092RV3	092RV3 rhizobium m
C 8	8	3.8	512 16	09HVA7	09HVA7 pseudomonas
C 9	8	3.8	791 17	09YB89	09YB89 aeropyrum p
C 10	8	3.8	1183 16	0980H9	0980H9 mycoplasma
C 11	8	3.8	3178 5	0969P4	0969P4 caenorhabd
C 12	8	3.3	55 10	041910	041910 arabidopsis
C 13	7	3.3	55 12	091FK1	091FK1 chilo iride
C 14	7	3.3	55 16	0920P8	0920P8 rhizobium m
C 15	7	3.3	65 11	09D860	09D860 mus musculu
C 16	7	3.3	65 11	09COP0	09COP0 mus musculu
C 17	7	3.3	72 6	09GMR4	09GMR4 mecatia fasc
C 18	7	3.3	85 5	09FCV5	09FCV5 lactobacilli
C 19	7	3.3	85 5	09VDM5	09VDM5 drosophila
C 20	7	3.3	91 8	099813	099813 ixodes hexa
C 21	7	3.3	105 16	09K7F6	09K7F6 bacillus ha
C 22	7	3.3	106 2	047966	047966 herpetosiph
C 23	7	3.3	117 8	099313	099313 nesticus st
C 24	7	3.3	117 8	099314	099314 nesticus st
C 25	7	3.3	117 8	099317	099317 nesticus ba
C 26	7	3.3	117 8	009333	009333 nesticus s1
C 27	7	3.3	120 12	057169	057169 vaccinia vl
C 28	7	3.3	121 2	09AL17	09AL17 enterococcu
C 29	7	3.3	123 17	0973P2	0973P2 sulfolobus
C 30	7	3.3	123 17	096ZM9	096ZM9 sulfolobus
C 31	7	3.3	124 12	P87608	P87608 compox viru
C 32	7	3.3	126 12	089067	089067 variola vir
C 33	7	3.3	172 5	095XJ8	095XJ8 caenorhabd
C 34	7	3.3	183 10	08RMP7	08RMP7 arabidopsi
C 35	7	3.3	191 12	09DVU7	09DVU7 plutella xy
C 36	7	3.3	195 17	059004	059004 pyrococcus
C 37	7	3.3	198 6	09GK23	09GK23 sus scrofa
C 38	7	3.3	201 16	08X6B2	08X6B2 escherichia
C 39	7	3.3	205 10	09SVW1	09SVW1 arabidopsi
C 40	7	3.3	207 5	022044	022044 caenorhabd
C 41	7	3.3	208 11	091ZW7	091ZW7 mus musculu
C 42	7	3.3	225 12	08VZJ1	08VZJ1 camelpox vl
C 43	7	3.3	228 5	09NSD7	09NSD7 caenorhabd
C 44	7	3.3	236 10	09AM49	09AM49 guillierdia
C 45	7	3.3	238 2	08VTA2	08VTA2 agrobacteri

ALIGNMENTS

RESULT 1
ID 097662 PRELIMINARY; PRT; 151 AA.
AC 097662;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Platelet activating factor receptor (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID:9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Sander F.C., Ibe O.B., Raj J.U.;
RT "Developmentally Regulated Expression of Platelet Activating Factor
Receptor in Ovine Lung."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF099674; AAC7459.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm.1; 1
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECPE_FL_1; UNKNOWN_1.

DR PROSITE: PS50262; G_PROTEIN_RECIP_FL_2; 1.
 KW Receptor.
 FT NON_TER 1
 FT NON_TER 151
 SO SEQUENCE 151 AA; 17001 MW; 31EC4F2006AF951C CRC64;

Alignment Scores:
 Pred. No.: 30.1 Length: 151
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q97662 (1-151)

OY 103 TCTTTTCTAGTTTCTAATA 80
 DB 138 SerpPheLeuValPheLeuLe 145

RESULT 2

Q9N5G4 PRELIMINARY; PRT; 280 AA.
 AC Q9N5G4: 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DF 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical 29.9 kDa protein.
 GN R05D8.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Feloderinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Langston Y., Wohlmann P., Lennox S.;
 RT "The sequence of C. elegans cosmid R05D8.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SBR) FAMILY

DR EMBL: AC006677; AAF39950.1; -
 DR HSSP: P50162; IAE1.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PRO0080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 KW Hypothetical protein; oxidoreductase.
 SO SEQUENCE 280 AA; 29923 MW; 58BC9AE6F1D5F006 CRC64;

Alignment Scores:
 Pred. No.: 26.6 Length: 280
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.77% Indels: 0
 DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q9N5G4 (1-280)

OY 516 AAAACTCTAAAGTTCAG 539
 DB 116 LysThreulysLeuasnLeuGln 123

RESULT 3

Q9GK76 PRELIMINARY; PRT; 342 AA.
 AC Q9GK76: 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DF 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Platelet-activating factor receptor.
 OS Capra hircus (Goat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Caprinae; Capra.
 OX NCBI_TaxId=9925;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang W., Diehl J.R., Piumi F.;
 RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor
 RT Gene.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF302764; AAG3982.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR01559; DUFYANTIGEN.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00337; G_PROTEIN_RECIP_FL_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECIP_FL_2; 1.
 KW Receptor.
 SO SEQUENCE 342 AA; 39612 MW; 39D9379AED60FD4E CRC64;

Alignment Scores:
 Pred. No.: 25.6 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: Gaps: 0

US-09-835-992A-20 (1-687) x Q9GK76 (1-342)

OY 103 TCTTTTCTAGTTTCTAATA 80
 DB 194 SerpPheLeuValPheLeuLe 201

RESULT 4

Q9TTY5 PRELIMINARY; PRT; 342 AA.
 AC Q9TTY5: 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DF 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Platelet-activating factor receptor.
 GN PAFR.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxId=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang W.S., Diehl J.R., Roudabush W.E.;
 RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor
 RT Gene.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
 Schmidt P.;
 RT "Molecular characterization of bovine platelet-activating factor

RT receptor transcripts and their detection in different tissues of cattle.
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF187321; AF01439.2; -
 DR EMBL: AJ295321; CAC43290.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR01559; DUFFYANTIGEN.
 DR PROSITE: PS00237; GPCR_RHODOPSIN.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
 DR PROSITE: PS00237; G_PROTEIN_RECEPTOR_FL2; 1.
 DR Receptor
 KM RECEPTOR
 SO SEQUENCE 342 AA; 39691 MW; 7C4236205AE937C9 CRC64;

Alignment Scores:
 Pred. No.: 25.6 Length: 342
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 6 Gaps: 0

US-09-835-992A-20 (1-687) x Q9TYS (1-342)
 QY 103 TCTTTTCTCTAGTTTCTAATA 80
 DB 194 Serpheeulvalphelulle 201
 RESULT 5
 Q17820 PRELIMINARY; PRT; 343 AA.
 AC Q17820;
 DT 01-NOV-1996 (TREMUREL. 01, Created)
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
 DE CO886.9 protein.
 GN CO886.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilkinson J.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99059613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C. elegans: A platform for investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z72502; CAA96591.1; -
 DR InterPro: IPR000594; Thif_domain.
 DR Pfam: PF00899; Thif; 1.
 DR SEQUENCE 343 AA; 38738 MW; 2897C0F918E881B6 CRC64;

Alignment Scores:
 Pred. No.: 25.6 Length: 343
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 5 Gaps: 0

US-09-835-992A-20 (1-687) x Q17820 (1-343)
 QY 628 CACATAAATAAGCACTAATA 605
 DB 23 Glnasnlytlearganserlys 30
 RESULT 6
 Q8TTZ5
 DR EMBL: AL591784; CAC45310.1; -

ID Q8TTZ5 PRELIMINARY; PRT; 352 AA.
 AC Q8TTZ5;
 DT 01-JUN-2002 (TREMUREL. 21, Created)
 DT 01-JUN-2002 (TREMUREL. 21, Last sequence update)
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
 DE Sulfate/molybdate ABC transporter, solute-binding protein.
 GN M0280.
 OS Methanobacteria acetylators.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nisbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Alnoor D., Brown A.,
 RA Allen N., Maylor J., Stange-Thomann N., DeArrellano K., Johnson R.,
 RA Linton L., McEwan P., McKernan K., Talamas J., Tjirelli A., Ye W.,
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
 RA Hedderich R., Ingram-Smith C., Kueltner H.C., Kraycek J.A.,
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Berry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.;
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
 RT and physiological diversity";
 RL Genome Res. 12:532-543(2002).
 DR EMBL: AE010687; AAM03733.1; -
 KW Complete proteome.
 SO SEQUENCE 352 AA; 38715 MW; 15652724987E8A9E CRC64;

Alignment Scores:
 Pred. No.: 25.4 Length: 352
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.79% Indels: 0
 DB: 17 Gaps: 0

US-09-835-992A-20 (1-687) x Q8TTZ5 (1-352)
 QY 100 TTTTCTCTAGTTTCTAATA 77
 DB 10 Phepheeulvalphelulle 17
 RESULT 7
 Q92RV3
 ID Q92RV3 PRELIMINARY; PRT; 376 AA.
 AC Q92RV3;
 DT 01-DEC-2001 (TREMUREL. 19, Created)
 DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
 DT 01-MAR-2002 (TREMUREL. 20, Last annotation update)
 DE Putative permease protein.
 GN R00738 OR SMC00793.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ransperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL: AL591784; CAC45310.1; -

DR InterPro: IPR000560; HisAc_phsphatse.
 DR InterPro: IPR002549; UPF0118.
 DR Pfam: PF01594; UPF0118; 1.
 DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 376 AA; 40673 MW; F119E2DFBF8CE87 CRC64;

Alignment Scores:

Pred. No.: 25.1 Length: 376
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.77% Indels: 0
 DB: 16 Gaps: 0

US-09-835-992A-20 (1-687) x Q9HVA7 (1-376)

OY 479 CNGTGGCATGTGTTCTTACTT 502

DB 97 LeuLeuAlaLeucysphLeuLeu 104

RESULT 8

O9HVA7 PRELIMINARY; PRT: 512 AA.

AC 09HVA7.
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Iron (III)-transport system permease HltB.
 GN HltB OR PA4688.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RP [1]

RC STRAIN-ATCC 15692 / PA01;

RA MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Eryin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";

RL Nature 406:959-964(2000).

CC -1- FUNCTION: PROBABLY PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT

CC SYSTEM, PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE

CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-

CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.

DR EMBL: AE004882; AAC08075.1; -;

DR InterPro: IPR000515; BPD_transp.

DR Pfam: PF00528; BPD_transp; 2.

DR PROSITE: PS00402; BPD_TRANS_PNN_MEMBR; 1.

KW Transmembrane; Transport; Complete proteome.

SQ SEQUENCE 512 AA; 56176 MW; 9A21A9B3CDEF0A CRC64;

Alignment Scores:

Pred. No.: 23.6 Length: 512
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.77% Indels: 0
 DB: 16 Gaps: 0

US-09-835-992A-20 (1-687) x Q9HVA7 (1-512)

OY 479 CNGTGGCATGTGTTCTTACTT 502

DB 230 LeuLeuAlaLeucysphLeuLeu 237

RESULT 9

O9YB89 PRELIMINARY; PRT: 791 AA.

AC 09YB89.
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein APEI708.

GN APEI708.

OS Aeropyrum pernix.

OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcates;

OC Desulfurococcaceae; Aeropyrum.

OX NCBI_TaxID=56636;

RP [1]

RC SEQUENCE FROM N.A.

RA STRAIN-K1.

RA MEDLINE=99310339; PubMed=10382966;

RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,

RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,

RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,

RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RT "Complete genome sequence of an aerobic hyper-thermophilic

RT crenarchaeon, Aeropyrum pernix K1.";

RL DNA Res. 6:83-101(1999).

DR EMBL: AP000062; BAA80709.1; -;

DR InterPro: IPR004089; Chmtaxis_transd.

DR InterPro: IPR002570; Hpt.

DR SMART: SM00073; HPT; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 791 AA; 85044 MW; D52098F863EBA82F CRC64;

Alignment Scores:

Pred. No.: 21.6 Length: 791
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.77% Indels: 0
 DB: 17 Gaps: 0

US-09-835-992A-20 (1-687) x Q9YB89 (1-791)

OY 76 TTGATATGAAAACTAGAAAAA 99

DB 134 LeuValLeuGluLysLeuLys 141

RESULT 10

O9BOH9 PRELIMINARY; PRT: 1183 AA.

AC 09BOH9.
 DT 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DE Hypothetical protein MYP0_3820.

GN MYP0_3820.

OS Mycoplasma pulmonis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

OC Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2107;

RP [1]

RC SEQUENCE FROM N.A.

RA STRAIN-UAB CTIP;

RA MEDLINE=21267165; PubMed=11353084;

RA Chambaud I., Helling R., Ferris S., Barbe V., Samson D., Gallisson F.,

RA Moszer I., Dydvig K., Wroblewski H., Vlati A., Kocha E.P.C.,

RA Blanchard A.;

RT "The complete genome sequence of the murine respiratory pathogen

RT Mycoplasma pulmonis.";

RL Nucleic Acids Res. 29:2145-2153(2001).

DR EMBL: AL445564; CAC13555.1; -;

DR MYP0_3820; -;

RA Darai G.;
RT "Identification and mapping of origins of DNA replication within the
RT DNA sequences of the genome of insect iridescent virus type 6.";
RL Virus Genes 6:19-32(1992).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE-93118242; PubMed-1475907;
RA Sonntag K.C., Darai G.;
RT "Characterization of the third origin of DNA replication of the genome
RT of insect iridescent virus type 6.";
RL Virus Genes 6:333-342(1992).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE-93260401; PubMed-8492091;
RA Stowasser R., Raab K., Schmitzler P., Janssen W., Darai G.;
RT "Identification of the gene encoding the major capsid protein of
RT insect iridescent virus type 6 by polymerase chain reaction.";
RL J. Gen. Virol. 74:873-879(1993).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE-94167241; PubMed-8121799;
RA Schmitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,
RT "Identification of genes encoding zinc finger proteins, non-histone
RT chromosomal HMG protein homologue, and a putative GTP phosphohydrolase
RT in the genome of Chilo iridescent virus.";
RL Nucleic Acids Res. 22:158-166(1994).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE-94353641; PubMed-8073636;
RA Sonntag K.C., Schmitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily. Implications for the
RT evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE-95213160; PubMed-7698884;
RA Sonntag K.C., Schmitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp)."
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE-94292906; PubMed-8021587;
RA Schmitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RT Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE-98141693; PubMed-9482589;
RA Bahr U., Tidon C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391; similarities in coding strategy between
RT insect and vertebrate Iridoviruses.";
RL Virus Genes 15:235-245(1997).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE-99125223; PubMed-9926400;
RA Muller K., Tidon C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE-99383793; PubMed-10456793;
RA Muller K., Tidon C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";

RL Virus Genes 18:243-264(1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE-21342589; PubMed-11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the first complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Irdescent Virus.";
RL Virology 286:182-196(2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF303741; AAK82184.1;
SQ SEQUENCE 55 AA; 6886 MW; 100D465DFB142CAE CRC64;

Alignment Scores:
Pred. No.: 411 Length: 55
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 16 Gaps: 0

US-09-835-992a-20 (1-687) x Q92UP8 (1-55)
Qy 529 AGTTAAGAGTTTCCTTGGG 509
Db 25 SerLeuArgValPheLeucly 31
RESULT 15
Q9D860

US-09-835-992a-20 (1-687) x Q91FK1 (1-55)
Qy 602 CCNAGAGACTATGTAACAA 582
Db 13 ProLysAsnTyrValAsnLys 19
RESULT 14
Q92UP8 PRELIMINARY; PRT; 55 AA.
AC Q92UP8;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical protein RB1047.
GN RB1047 OR Smb21672.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymB (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE-21396508; PubMed-11481431;
RA Flinn-T.M., Weidner S., Wong K., Buhmester J., Chain P.,
RT Vornheller F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puhler A.;
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603645; CAC94447.1;
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 55 AA; 6547 MW; E20E5E35361C1CB CRC64;

Alignment Scores:
Pred. No.: 411 Length: 55
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.32% Indels: 0
DB: 16 Gaps: 0

US-09-835-992a-20 (1-687) x Q92UP8 (1-55)
Qy 529 AGTTAAGAGTTTCCTTGGG 509
Db 25 SerLeuArgValPheLeucly 31
RESULT 15
Q9D860

ID Q9D860 PRELIMINARY; PRT; 65 AA.
AC Q9D860;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 0610009M10R1k protein.
GN MRL33 OR 0610009M10R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Araiawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Salto T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK008431; BAB25665.1;
DR MGD; MGI:2137225; MRP133.
SQ SEQUENCE 65 AA; 7344 MW; 35BE69CA0359BE6C CRC64;

Alignment Scores:

Prod. No.:	397	Length:	65
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.32%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992a-20 (1-687) x Q9D860 (1-65)

OY 590 GTAAACAAAAAGTTTGT 570

|||||

DB 50 VALASNLVSLVSVALLAUPHE 56

Search completed: January 14, 2003, 17:36:55
Job time : 54.3743 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 ; Search time 48.9632 Seconds
(without alignments)

5410.226 Million cell updates/sec

Title: US-09-835-992a-21

Sequence: 1 ctccaccagctgcctccag.....ccaccctgggtgtgtggat 994

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame_n2p_model -DEV-x1p
-O-/c992.1/USPTO.spool/US09835992/runal.14012003.161525.3979/app-query.fasta.1.3932
-DB-A_Geneseq.101002 -OPMT-fastan -SUFFIX-oligo -TAG -MINMATCH-0.1 -LOOPEL-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi
-LIST-45 -DOCALLIN-200 -THR.SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFM-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USR-US09835992.c992.1.121.ernal.14012003.161525.3979 -NCPU-6 -ICPU-3
-NO.XLXY -NO.MAP -LARGEQUERY -NMG_SCORES-0 -WAIT -LONGLOG -DEV-TIMEOUT-120
-MARN-TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database :

A_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	50	15.9	547	23	ABB57301	Mouse ischaemic co
2	30	9.6	378	22	ABB65147	Drosophila melanog
3	30	9.6	544	22	ABB65056	Drosophila melanog
4	9	2.9	397	22	AAE09042	Equine influenza v
5	9	2.9	716	22	AAE09044	Equine influenza v
6	8	2.5	153	21	AAE02038	Arabidopsis thalia
7	8	2.5	182	21	AAE09091	Arabidopsis thalia
8	8	2.5	235	21	AAE020137	Arabidopsis thalia
9	8	2.5	241	21	AAE020136	Arabidopsis thalia
10	8	2.5	241	21	AAE020137	Arabidopsis thalia
11	8	2.5	248	21	AAE039090	Arabidopsis thalia
12	8	2.5	261	21	AAE037797	Arabidopsis thalia
13	8	2.5	284	21	AAE039054	Arabidopsis thalia
14	8	2.5	291	21	AAE031127	Arabidopsis thalia
15	8	2.5	291	21	AAE037795	Arabidopsis thalia
16	8	2.5	295	21	AAE030200	Arabidopsis thalia
17	8	2.5	295	21	AAE044684	Zea mays protein f
18	8	2.5	301	21	AAE039089	Arabidopsis thalia
19	8	2.6	302	21	AAE022946	Arabidopsis thalia
20	8	2.6	306	21	AAE022945	Arabidopsis thalia
21	8	2.6	314	21	AAE022944	Arabidopsis thalia
22	8	2.6	314	21	AAE039926	Arabidopsis thalia
23	8	2.6	314	23	ABE03213	Herbicidally activ
24	8	2.5	317	21	AAE036745	Arabidopsis thalia
25	8	2.6	325	21	AAE039925	Arabidopsis thalia
26	8	2.5	330	21	AAE039053	Arabidopsis thalia
27	8	2.5	333	21	AAE039924	Arabidopsis thalia
28	8	2.5	334	21	AAE044646	Arabidopsis thalia
29	8	2.5	345	21	AAE01126	Arabidopsis thalia
30	8	2.5	345	21	AAE030189	Arabidopsis thalia
31	8	2.5	345	21	AAE044683	Arabidopsis thalia
32	8	2.5	383	21	AAE036744	Zea mays protein f
33	8	2.5	384	22	AAE024296	Regulatory protein
34	8	2.5	397	22	AAE036868	Staphylococcus aur
35	8	2.5	411	21	AAE039052	Arabidopsis thalia
36	8	2.5	414	21	AAE01125	Arabidopsis thalia
37	8	2.5	414	21	AAE039458	Arabidopsis thalia
38	8	2.5	427	22	AAE020010	Arabidopsis thalia
39	8	2.5	436	21	AAE036743	Brassica 3-ketocacy
40	8	2.5	436	23	ABE02008	Herbicidally activ
41	8	2.5	443	21	AAE044645	Arabidopsis thalia
42	8	2.5	443	22	AAE020005	Arabidopsis 3-keto
43	8	2.5	446	21	AAE044644	Arabidopsis thalia
44	8	2.5	453	21	AAE039457	Arabidopsis thalia
45	8	2.5	457	21	AAE039456	Arabidopsis thalia

ALIGNMENTS

RESULT 1
/ABB57301

ID ABB57301 standard; Protein: 547 AA.

XX ABB57301;

XX 07-MAR-2002 (first entry)

XX Mouse ischaemic condition related protein sequence SEQ ID NO:843.

XX Mouse: ischaemia; compressive ischaemia; occlusive ischaemia;

XX vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

XX WO20018188-A2.

XX 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.
 XX
 PR 18-MAY-2000; 2000JP-0145977.
 XX
 XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 PA
 XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 PI
 XX WPI; 2002-034733/04.
 DR N-PSDB; ABI99762.
 XX
 PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these
 PT genes -
 XX
 PS Claim 2; Page 2091-2093; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (1) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (1). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (ABI99202 to ABI99912, encoding
 CC the protein sequences in AB57020 to AB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.
 XX
 SQ Sequence 547 AA;
 Alignment Scores:
 Pred. No.: 7.87e-44 Length: 547
 Score: 50.00 Matches: 50
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.92% Indels: 0
 DB: 23 Gaps: 0
 US-09-835-992A-21 (1-994) x AB57301 (1-547)
 QY 522 ACCTGTTGATAGAGAGATATACATATGAGAAAGTGGCTCAATCCTAGTGGT 581
 DB 329 ThrLeuValAspArgGlyAspAsnThrTyrglyGlyTyrPallleAsnProSerGly 348
 QY 582 GGACTGATTTCAAAGGACACCCACTAGAGCGCTACAGGTCTGCTCAGTGTGCAACTC 641
 DB 349 GlyLeuIleSerGlyGlyHisProLeuGlyAlaThrGlyLeuAlaGlyCysAlaGluLeu 368
 QY 642 TGCTGCAGCTGAGAGGGAAGCCGAAAA 671
 DB 369 CysTrpGlnLeuArgGlyGlnAlaGlyLys 378
 RESULT 2
 ABB65147
 ID ABB65147 standard; Protein; 378 AA.
 XX
 AC ABB65147;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 22233.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PF

XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 PF
 XX 23-MAR-2000; 2000US-191637P.
 PR
 XX 11-JUL-2000; 2000US-0614150.
 PR
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EM;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL09250.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 22233; 21pp + Sequence listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB116175) and the encoded proteins
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 378 AA;
 Alignment Scores:
 Pred. No.: 1.22e-22 Length: 378
 Score: 30.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.55% Indels: 0
 DB: 22 Gaps: 0
 US-09-835-992A-21 (1-994) x ABB65147 (1-378)
 QY 570 AATCCTAGTGGTGGACTGATTTCAAAGGACACCCACTAGGCGCTCAGTCTGCTCAG 629
 DB 308 AsnProSerGlyGlyLeuIleSerGlyHisProLeuGlyAlaThrGlyLeuAlaGln 327
 QY 630 TGTGCAGAACTCTGCTGCGCAGCTGAGAGG 659
 DB 328 CysAlaGlnLeuGlyTrpGlnLeuArgGly 337
 RESULT 3
 ABB65056
 ID ABB65056 standard; Protein; 544 AA.
 XX
 AC ABB65056;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 21960.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, LI PWD, Myers EW:
XX WPI; 2001-656860/75.
DR N-PSDB; ABL09159.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 21960; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 544 AA;
XX
Alignment Scores:
Pred. No.: 1.18e-22 Length: 544
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.55% Indels: 0
DB: 22 Gaps: 0
XX
US-09-835-992a-21 (1-994) x ABB65056 (1-544)
XX
QY 570 AATCTGATGGTGGACGATTTCAAGGACACCTAGCGCTGCTGCTGAG 629
DB 337 AabProserelgylgylleuileserlglh1eProleuglylalnlglyleuAla61n 356
XX
QY 630 TGTGACAGACTGTGCTGGACGACGAGGG 659
DB 357 CysAla61leuency8TTPG1nleuA9rgly 366
XX
RESULT 4
AAE09042
ID AAE09042 standard; Protein; 397 AA.
XX
AC AAE09042;
XX
DT 15-NOV-2001 (first entry)
XX
DE Equine Influenza virus H3N8 PelcapA-N-397 protein.
XX
KW Equine Influenza virus; ei: cold adaptation; temperature sensitivity;
KW vaccine; pelcapA-N-1217 DNA; PelcapA-N-397 protein.
XX
OS Equine Influenza virus H3N8.
XX
PN WO200160849-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US05048.
XX
PR 16-FEB-2000; 2000US-0506286.
XX
PA (UYPI-) UNIV PITTSBURGH.

XX
PI Dowling PW, Youngner JS;
XX WPI; 2001-522584/57.
DR N-PSDB; AAD15714.
XX
PT Novel isolated equine Influenza virus (wild-type and cold-adapted)
PT proteins and viruses containing nucleic acid molecules encoding the
PT proteins, which are useful for protecting animals from Influenza virus
PT infections -
XX
PS Claim 5; Page 137-138; 172pp; English.
XX
CC The patent discloses cold-adapted equine Influenza viruses and
CC reassortant Influenza A viruses comprising at least one genome
CC segment of such an equine Influenza virus, wherein the equine
CC Influenza virus genome segment confers at least one identifying
CC phenotype of the cold-adapted equine Influenza virus, such as
CC cold adaptation, temperature sensitivity, dominant interference
CC or attenuation. The viruses are useful for protecting animals
CC from diseases caused by Influenza viruses. They are also used
CC as vaccines. The present sequence is equine Influenza (ei) virus
CC H3N8 Pelca (cold adapted) PA-N-397 protein which is encoded by
CC pelcapA-N-1217 DNA.
XX
SQ Sequence 397 AA;
XX
Alignment Scores:
Pred. No.: 2.07 Length: 397
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.90% Indels: 0
DB: 22 Gaps: 0
XX
US-09-835-992a-21 (1-994) x AAE09042 (1-397)
XX
QY 337 AACGAGCTTGCAATCATGATCATTT 311
DB 55 AsnGluLeuEnGlyLysSerValIleIle 63
XX
RESULT 5
AAE09044
ID AAE09044 standard; Protein; 716 AA.
XX
AC AAE09044;
XX
DT 15-NOV-2001 (first entry)
XX
DE Equine Influenza virus H3N8 PelcapA-716 protein.
XX
KW Equine Influenza virus; ei: cold adaptation; temperature sensitivity;
KW vaccine; pelcapA-2148 DNA; PelcapA-716 protein.
XX
OS Equine Influenza virus H3N8.
XX
PN WO200160849-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US05048.
XX
PR 16-FEB-2000; 2000US-0506286.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Dowling PW, Youngner JS;
XX WPI; 2001-522584/57.
DR N-PSDB; AAD15716.
XX
PT Novel isolated equine Influenza virus (wild-type and cold-adapted)
PT proteins and viruses containing nucleic acid molecules encoding the

PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156658.
PR 29-SEP-1999; 99US-0156696.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158993.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 26.2
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Caps: 0

US-09-835-992a-21 (1-994) x AAG20138 (1-153)
QY 597 GCACACCCACTAGCGCTACAGT 620
Db 89 GlyHisProLeuGlyAlaThrGly 96

RESULT 7
AAG39091
ID AAG39091 standard; Protein; 182 AA.
XX
AC AAG39091;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48319.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridization assay; genetic mapping; gene expression control; promoter;
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130049.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

[illegible]

PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161922.
PR 28-OCT-1999: 99US-0161933.
PR 29-OCT-1999: 99US-0162142.

Alignment Scores:
Pred. No.: 25.7 Length: 182
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-21 (1-994) x AAG39091 (1-182)
Oy 67 MAATGACACTTGCMAAATTG 90
Db 138 LysleuAsnThrleuGlnLysleu 145
|||||
RESULT 8
AAG20137
ID AAG20137 standard; Protein; 187 AA.
XX AAG20137;
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 22207.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 200DEP-0301439.
XX
XX
PR 25-FEB-1999: 99US-0121825.
PR 05-MAR-1999: 99US-0123180.
PR 09-MAR-1999: 99US-0123548.
PR 23-MAR-1999: 99US-0125788.
PR 25-MAR-1999: 99US-0126264.
PR 29-MAR-1999: 99US-0126785.
PR 01-APR-1999: 99US-0127462.
PR 06-APR-1999: 99US-0128234.
PR 08-APR-1999: 99US-0128714.
PR 16-APR-1999: 99US-0129845.
PR 19-APR-1999: 99US-0130077.
PR 21-APR-1999: 99US-0130449.
PR 23-APR-1999: 99US-0130510.
PR 23-APR-1999: 99US-0130891.
PR 28-APR-1999: 99US-0131449.
PR 30-APR-1999: 99US-0132048.
PR 30-APR-1999: 99US-0132407.
PR 04-MAY-1999: 99US-0132484.
PR 05-MAY-1999: 99US-0132485.
PR 06-MAY-1999: 99US-0132486.
PR 06-MAY-1999: 99US-0132487.
PR 07-MAY-1999: 99US-0132863.

PR 11-MAY-1999: 99US-0134256.
PR 14-MAY-1999: 99US-0134218.
PR 14-MAY-1999: 99US-0134219.
PR 14-MAY-1999: 99US-0134321.
PR 14-MAY-1999: 99US-0134370.
PR 18-MAY-1999: 99US-0134768.
PR 19-MAY-1999: 99US-0134941.
PR 20-MAY-1999: 99US-0135124.
PR 21-MAY-1999: 99US-0135353.
PR 24-MAY-1999: 99US-0135629.
PR 25-MAY-1999: 99US-0136021.
PR 27-MAY-1999: 99US-0136392.
PR 28-MAY-1999: 99US-0136782.
PR 01-JUN-1999: 99US-0137222.
PR 03-JUN-1999: 99US-0137528.
PR 04-JUN-1999: 99US-0137502.
PR 07-JUN-1999: 99US-0137724.
PR 08-JUN-1999: 99US-0138094.
PR 10-JUN-1999: 99US-0138540.
PR 10-JUN-1999: 99US-0138847.
PR 14-JUN-1999: 99US-0139119.
PR 16-JUN-1999: 99US-0139452.
PR 16-JUN-1999: 99US-0139453.
PR 17-JUN-1999: 99US-0139492.
PR 18-JUN-1999: 99US-0139454.
PR 18-JUN-1999: 99US-0139455.
PR 18-JUN-1999: 99US-0139456.
PR 18-JUN-1999: 99US-0139457.
PR 18-JUN-1999: 99US-0139458.
PR 18-JUN-1999: 99US-0139459.
PR 18-JUN-1999: 99US-0139460.
PR 18-JUN-1999: 99US-0139461.
PR 18-JUN-1999: 99US-0139462.
PR 18-JUN-1999: 99US-0139463.
PR 18-JUN-1999: 99US-0139750.
PR 18-JUN-1999: 99US-0139763.
PR 21-JUN-1999: 99US-0139817.
PR 22-JUN-1999: 99US-0139899.
PR 23-JUN-1999: 99US-0140353.
PR 23-JUN-1999: 99US-0140354.
PR 24-JUN-1999: 99US-0140695.
PR 28-JUN-1999: 99US-0140823.
PR 29-JUN-1999: 99US-0140991.
PR 30-JUN-1999: 99US-0141287.
PR 01-JUL-1999: 99US-0141842.
PR 01-JUL-1999: 99US-0142154.
PR 02-JUL-1999: 99US-0142055.
PR 06-JUL-1999: 99US-0142390.
PR 08-JUL-1999: 99US-0142803.
PR 09-JUL-1999: 99US-0142920.
PR 12-JUL-1999: 99US-0142977.
PR 13-JUL-1999: 99US-0143542.
PR 14-JUL-1999: 99US-0143624.
PR 15-JUL-1999: 99US-0144005.
PR 16-JUL-1999: 99US-0144085.
PR 16-JUL-1999: 99US-0144086.
PR 19-JUL-1999: 99US-0144325.
PR 19-JUL-1999: 99US-0144331.
PR 19-JUL-1999: 99US-0144332.
PR 19-JUL-1999: 99US-0144333.
PR 19-JUL-1999: 99US-0144334.
PR 19-JUL-1999: 99US-0144335.
PR 20-JUL-1999: 99US-0144352.
PR 20-JUL-1999: 99US-0144632.
PR 20-JUL-1999: 99US-0144884.
PR 21-JUL-1999: 99US-0144814.
PR 21-JUL-1999: 99US-0145086.
PR 21-JUL-1999: 99US-0145088.
PR 22-JUL-1999: 99US-0145085.
PR 22-JUL-1999: 99US-0145087.
PR 22-JUL-1999: 99US-0145089.
PR 22-JUL-1999: 99US-0145192.
PR 23-JUL-1999: 99US-0145145.

```

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0145386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

```

```

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 25.6
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.5%
DB: 21
Gaps: 0

US-09-835-992A-21 (1-994) x AAG20137 (1-187)
QY 597 GCACACCCACTAGCGCTACAGT 620
DB 123 GLYHISProLeuGLyAlaThrGLy 130

RESULT 9
AAG20136
ID AAG20136 standard; Protein; 235 AA.
XX
AC AAG20136;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22206.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

```

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158202.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160960.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Alignment Scores:
Pred. No.: 25
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Gaps: 0
Length: 235
Matches: 8
Conservative: 0
Mismatches: 0
Indels: 0
US-09-835-992a-21 (1-994) x AAG20136 (1-235)
OY 597 GGACACCCACTAGCGCTACAGCT 620
DB 171 GlyHisProLeuGlyAlaThrGly 178
RESULT 10
AAG37797
ID AAG37797 standard; Protein; 241 AA.
XX AAG37797;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46533.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139869.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
pred. No.: 25
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Gaps: 0
Length: 241
Matches: 8
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-835-992a-21 (1-994) x AAG37797 (1-241)
QY 597 GGACACCACTAGCGCTACAGGT 620
Db 177 GYHISProteuclYalathrgly 184
RESULT 11
AAG39090
ID AAG39090 standard; Protein; 248 AA.
AC AAG39090;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48318.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134421.
PR 14-MAY-1999; 99US-0134570.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-01394941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144881.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147492.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

```
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Prod. No.: 24, 9
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21
Gaps: 0

US-09-835-992a-21 (1-994) x AAG39090 (1-248)
OY 67 AAGTGAACACTTGCAGAAATTC 90
Db 204 LysleuAnthrLeuGlnLysLeu 211

RESULT 12
AAG37796
ID AAG37796 standard; Protein: 261 AA.
AC AAG37796;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 46532.
KW Protein identification: signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121625.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0144624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 23-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
```

PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157111.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 24.8
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.55%
DB: 21

US-09-835-992a-21 (1-994) x AAG37796 (1-261)
OY 597 GGACACCACTAGCGCTACAGT 620
DB 197 GLYHISPROLENGLYALATHRGly 204

RESULT 13
AAG39054
ID AAG39054 standard; Protein; 284 AA.
XX
AC AAG39054;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48269.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121801.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0138840.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142290.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142927.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143524.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Alignment Scores:
Pred. No.: 24.6 Length: 284
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
US-09-835-992a-21 (1-994) x AAG39054 (1-284)
QY 597 GGACACCCACTAGCGCTACAGCT 620
DB 225 GLYHISProLeuglyAlaThrly 232
RESULT 14
ID AAG11127 standard; Protein; 291 AA.
XX AAG11127;
AC AAG11127;
XX 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 9727.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154479.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Wed Jan 15 08:41:12 2003

Pred. No.:	24.5	Length:	291
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992a-21 (1-994) x ANG37795 (1-291)

OY 597 CGACACCCACTAGGCGCTACAGCT 620
|||||

Db 227 GlyHleProLeuGlyAlaThrGly 234

Search completed: January 14, 2003, 17:27:24
Job time : 53.9632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 : Search time 13.6606 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992a-21

Perfect score: 314
Sequence: 1 ctcaccagctgctcctcag.....ccaccctgggtggtggat 994

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-O=/cgn2_1/uspro.spool/US0983599/runat.14012003.161527.4050/app_query.fasta.1.3932
-DB=Issued_Patents_AA_QPMT-fastn -SUFFIX=olig.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992.ecgn.1.1.49.8runat.14012003.161527.4050 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT-120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	8	2.5	2	US-08-286-819A-40 Sequence 40, Appl
2	8	2.5	259	US-08-980-357-40 Sequence 40, Appl
3	8	2.5	384	US-08-286-819A-14 Sequence 14, Appl
4	8	2.5	384	US-08-980-357-14 Sequence 14, Appl
5	8	2.5	2254	US-08-286-819A-28 Sequence 28, Appl
6	8	2.5	2254	US-08-980-357-28 Sequence 28, Appl
7	7	2.2	15	US-09-045-632-103 Sequence 103, Appl
8	7	2.3	3	PCT-US95-12080-4 Sequence 4, Appl
9	7	2.2	119	US-08-983-607-34 Sequence 34, Appl
10	7	2.3	188	US-08-737-109-14 Sequence 14, Appl
11	7	2.2	219	US-08-855-140-4 Sequence 4, Appl
12	7	2.3	260	US-08-081-929-10 Sequence 10, Appl

13	7	2.2	268	3	US-08-071-483-11	Sequence 11, Appl
14	7	2.3	270	1	US-08-452-052-1	Sequence 1, Appl
15	7	2.3	274	4	US-08-936-165A-499	Sequence 499, App
16	7	2.3	290	4	US-09-230-637-32	Sequence 32, Appl
17	7	2.2	291	4	US-09-134-001C-5669	Sequence 5669, Ap
18	7	2.3	298	3	US-09-025-691-1	Sequence 1, Appl
19	7	2.3	345	1	US-08-134-570-14	Sequence 14, Appl
20	7	2.3	373	1	US-08-015-986A-14	Sequence 14, Appl
21	7	2.3	373	2	US-08-446-163-14	Sequence 14, Appl
22	7	2.2	426	4	US-08-961-083-48	Sequence 48, Appl
23	7	2.2	427	3	US-09-196-857-2	Sequence 2, Appl
24	7	2.2	458	4	US-09-504-358-2	Sequence 2, Appl
25	7	2.2	458	4	US-09-504-314-2	Sequence 2, Appl
26	7	2.3	471	4	US-09-391-104-25	Sequence 25, Appl
27	7	2.3	505	1	US-08-221-750A-5	Sequence 5, Appl
28	7	2.2	540	3	US-08-964-268-6	Sequence 6, Appl
29	7	2.2	545	2	US-08-426-125-2	Sequence 2, Appl
30	7	2.2	585	2	US-08-455-355-2	Sequence 2, Appl
31	7	2.2	610	4	US-09-443-184-55	Sequence 55, Appl
32	7	2.3	785	4	US-08-374-077C-3	Sequence 3, Appl
33	7	2.3	785	4	US-08-895-590-3	Sequence 3, Appl
34	7	2.3	785	4	US-09-539-879A-3	Sequence 3, Appl
35	7	2.2	820	1	US-08-291-896-2	Sequence 2, Appl
36	7	2.2	820	2	US-08-485-278-2	Sequence 2, Appl
37	7	2.2	906	1	US-08-254-573-2	Sequence 2, Appl
38	7	2.2	906	1	US-08-687-379-2	Sequence 2, Appl
39	7	2.2	906	1	US-08-687-379-4	Sequence 4, Appl
40	7	2.2	906	4	US-08-172-322-1	Sequence 1, Appl
41	7	2.2	906	4	US-08-216-326-2	Sequence 2, Appl
42	7	2.2	907	1	US-07-718-575-2	Sequence 2, Appl
43	7	2.2	907	1	US-08-481-206-2	Sequence 2, Appl
44	7	2.2	907	2	US-08-486-269A-2	Sequence 2, Appl
45	7	2.2	1286	6	5206163-1	Patent No. 5206163

ALIGNMENTS

RESULT 1
US-08-286-819A-40
Sequence 40, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEIDS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA: FR 9013579
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: E. faecium
STRAIN: BM4147
US-08-286-819A-40

Alignment Scores:
Pred. No.: 7.98 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: Gaps: 0

US-09-835-992a-21 (1-994) x US-08-286-819A-40 (1-259)

OY 58 ATGAAACAATAATGACACTTG 81
|||||
DB 5 MetGluGlnLysLeuAsnThrIeu 12

RESULT 2
US-08-980-357-40

Sequence 40, Application US/08980357
Patent No. 6013508

GENERAL INFORMATION:

APPLICANT: ARTHUR, MICHEL

APPLICANT: DUKTA-MALEN, SYLVIE

APPLICANT: MOLINAS, CATHERINE

APPLICANT: COURVALIN, PATRICE

TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE

TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR

TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357

FILING DATE: 435
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 259 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: E. faecium
STRAIN: BM4147
US-08-980-357-40

Alignment Scores:
Pred. No.: 7.98 Length: 259
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
DB: Gaps: 0

US-09-835-992a-21 (1-994) x US-08-980-357-40 (1-259)

OY 58 ATGAAACAATAATGACACTTG 81
|||||

DB 5 MetGluGlnLysLeuAsnThrIeu 12

RESULT 3
US-08-286-819A-14

Sequence 14, Application US/08286819A
Patent No. 5871910

GENERAL INFORMATION:

APPLICANT: ARTHUR, MICHEL

APPLICANT: DUKTA-MALEN, SYLVIE

APPLICANT: MOLINAS, CATHERINE

APPLICANT: COURVALIN, PATRICE

TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE

TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR

TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR

NUMBER OF SEQUENCES: 54

CORRESPONDENCE ADDRESSES:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-14

Alignment Scores:
Pred. No.: 7.63 Length: 384
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 2

US-09-835-992A-21 (1-994) x US-08-286-819A-14 (1-384)

OY 58 ATGGAACAAATTTGAACACTTG 81
|||||
Db 130 MetGlugInLysLeuAsnThrLeu 137

RESULT 4
US-08-980-357-14
Sequence 14, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEIDS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia

COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-14

Alignment Scores:
Pred. No.: 7.63 Length: 384
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 3

US-09-835-992A-21 (1-994) x US-08-980-357-14 (1-384)

OY 58 ATGGAACAAATTTGAACACTTG 81
|||||
Db 130 MetGlugInLysLeuAsnThrLeu 137

RESULT 5
US-08-286-819A-28
Sequence 28, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEIDS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400

CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAF UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-28

Alignment Scores:
Pred. No.: 6.27 Length: 2254
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.538 Indels: 0
Gaps: 0
DB: 2

US-09-835-992a-21 (1-994) x US-08-286-819A-28 (1-2254)

OY 58 ATGGAACAAATGACACTTG 81
|||||

DB 808 MetGlunLysLeuasnThrleu 815

RESULT 6
US-08-980-357-28
Sequence 28, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: CORVALAN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMERES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:

ADDRESSER: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAF UR
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 2254 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-980-357-28

Alignment Scores:
Pred. No.: 6.27 Length: 2254
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 3

US-09-835-992a-21 (1-994) x US-08-980-357-28 (1-2254)

OY 58 ATGGAACAAATGACACTTG 81
|||||

DB 808 MetGlunLysLeuasnThrleu 815

RESULT 7
US-09-045-632-103
Sequence 103, Application US/09045632
Patent No. 6001575
GENERAL INFORMATION:
APPLICANT: Huganir, Richard L.
APPLICANT: Dong, Hualing
TITLE OF INVENTION: THERAPEUTIC USES OF GRIP AND
TITLE OF INVENTION: GRIP-RELATED MOLECULES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/045,632
FILING DATE: 19-MAR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,016
FILING DATE: 19-MAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48147/1699-CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-045-632-103

Alignment Scores:
Pred. No.: 115 Length: 15
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
Gaps: 0
DB: 3

US-09-835-992a-21 (1-994) x US-09-045-632-103 (1-15)

QY 603 CCACCTAGCCGCTACAGCTCTT 623
Db 9 ProlenGlyAlaIthrGlyLeu 15

RESULT 8
PCT-US95-12080-4
Sequence 4, Application PC/TUS9512080
GENERAL INFORMATION:
APPLICANT: Children's Medical Center Corporation
TITLE OF INVENTION: Synducin Mediated Modulation of Tissue Repair
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/12080
FILING DATE:
CLASSIFICATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-815-8795

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-12080-4

Alignment Scores:
Pred. No.: 105 Length: 33
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
Gaps: 0
DB: 5

US-09-835-992a-21 (1-994) x PCT-US95-12080-4 (1-33)

QY 194 CAAAAATCAAAACTTCTTT 174
Db 19 GlnIlyIleIysAsnIlePhe 25

RESULT 9
US-08-983-607-34
Sequence 34, Application US/08983607
Patent No. 6140470
GENERAL INFORMATION:
APPLICANT: Alan Garen
TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
TITLE OF INVENTION: bodies
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Department of Molecular Biophysics
and Biochemistry, Yale University
STREET: 266 Whitney Avenue
CITY: New Haven
STATE: Connecticut
COUNTRY: United States of America
ZIP: 06520-8114
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 MB diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processing
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,607
FILING DATE: April 27, 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IB96/01032
FILING DATE: June 28, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Mary M. Krinsky
REGISTRATION NUMBER: 32423
REFERENCE/DOCKET NUMBER: OCR-679
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-773-9544
TELEFAX: 203-773-1183
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
DESCRIPTION:
ORIGINAL SOURCE:
ORGANISM: Homo sapiens (melanoma patient immu-
ORGANISM: nized with autologous tumor cells)
INDIVIDUAL ISOLATE: peripheral blood lymphocytes
IMMEDIATE SOURCE:
LIBRARY: DMA14 scfv antibodies obtained from

LIBRARY: fuses fusion phage construct
CLONE: V373
FEATURE:
NAME/KEY: heavy chain
OTHER INFORMATION: Xaa at position 3 is Gln or
OTHER INFORMATION: Xaa and Xaa at position 19 is Xaa or Ser
US-08-983-607-34

Alignment Scores:
Pred. No.: 91 Length: 119
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-21 (1-994) x US-08-983-607-34 (1-119)

OY 564 CTCATTAATCCTAGTGCTGA 584
DB 48 VALLIeasnpIroserGly 54

RESULT 10
US-08-737-109-14
Sequence 14, Application US/08737109
Patent No. 643568
GENERAL INFORMATION:
APPLICANT: SLABAS, Antoni Ryszard
APPLICANT: ELBOROUGH, Kieran Michael
APPLICANT: BRIGHT, Simon William Jonathan
APPLICANT: FENTIM, Philip Anthony
TITLE OF INVENTION: Plant Gene Specifying Acetyl Coenzyme A
NUMBER OF SEQUENCES: 32
TITLE OF INVENTION: Carboxylase and Transformed Plants Containing Same
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,109
FILING DATE: 21-OCT-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/00846
FILING DATE: 02-MAY-1994
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 188 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: YES
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE: Arabidopsis thaliana
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pKLU81
US-08-737-109-14

Alignment Scores:
Pred. No.: 86.5 Length: 188
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.25% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-21 (1-994) x US-08-737-109-14 (1-188)

OY 269 CTCCTGACAAATGCTTACT 249
DB 154 LeuLeuTyrLysCysPheThr 160

RESULT 11
US-08-855-140-4
Sequence 4, Application US/08855140
Patent No. 5854022
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Guegler, Karl J.
TITLE OF INVENTION: NEW TRANSMEMBRANE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/855,140
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0296 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1279546
US-08-855-140-4

Alignment Scores:
Pred. No.: 85 Length: 219
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-21 (1-994) x US-08-855-140-4 (1-219)

OY 61 GAACAAAATTGACACCTTG 81
DB 107 GluGlnLysLeuAsnThrLeu 113

RESULT 12
US-08-081-929-10
Sequence 10, Application US/08081929

Patent No. 6160090
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Barnea, Gilad
APPLICANT: Grunet, Martin H.
APPLICANT: Margolis, Richard U.
TITLE OF INVENTION: A NEW CLASS OF RPPASES: THEIR
STRUCTURAL DOMAINS AND LIGANDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,929
FILING DATE: 23-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30742
REFERENCE/DOCKET NUMBER: 7683-041-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-081-929-10

Alignment Scores:
Pred. No.: 83.4 Length: 260
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-21 (1-994) x US-08-081-929-10 (1-260)

OY 942 CTTGAGCGCAAAATTCN 922
DB 114 Lcugluclyglnlysphepro 120

RESULT 13
US-08-871-483-11
Sequence 11, Application US/08871483
GENERAL INFORMATION:
APPLICANT: Georgiou, George
APPLICANT: Qui, Ji
APPLICANT: Besette, Paul
APPLICANT: Swartz, James
TITLE OF INVENTION: METHODS FOR PRODUCING HETEROLOGOUS
NUMBER OF SEQUENCES: 13
DISULFIDE BOND-CONTAINING PEPTIDES IN BACTERIAL CELLS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas

COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,483
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: UTSH:614
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-871-483-11

Alignment Scores:
Pred. No.: 83.1 Length: 268
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 3 Gaps: 0

US-09-835-992a-21 (1-994) x US-08-871-483-11 (1-268)

OY 231 GCAGCAGCATTTGGCCAGT 251
DB 183 AlaAlaAlaIleuAlaSer 189

RESULT 14
US-08-452-052-1
Sequence 1, Application US/08452052
GENERAL INFORMATION:
APPLICANT: Poles, Elliot
TITLE OF INVENTION: FUNCTIONAL LIGANDS FOR THE AXONAL CELL
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,052
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-075
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 270 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-452-052-1

Alignment Scores:
Pred. No.: 83.1 Length: 270
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-452-052-1 (1-270)

QY 942 CTTGAAGCGCAAAATTCCTCCN 922
|||||
DB 114 LeuGlulGlyInlYsPhePro 120

RESULT 15
US-08-936-165A-499
Sequence 499, Application US/08936165A
Patent No. 6348582

GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 499:
SEQUENCE CHARACTERISTICS:
LENGTH: 274 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-499

Alignment Scores:
Pred. No.: 82.9 Length: 274
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-21 (1-994) x US-08-936-165A-499 (1-274)

QY 773 GC6GCTTCGCGGAACCCCA 753
|||||
DB 137 AlaAlaSerGlyIysThrPro 143

Search completed: January 14, 2003, 17:42:16
Job time: 20.9106 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 9.51637 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-21
Page: 214

Sequence: 1 ctcaaccagttgctcctcag.....ccaacctgggtgtgtgat 994

Xgapop	60.0	,	Xgapext	60.0
Ygapop	60.0	,	Ygapext	60.0
Fgapop	6.0	,	Fgapext	7.0
Delop	6.0	,	Delext	7.0

Searched: 118974 seqs, 19401057 residues

Word size:

Total number of hits satisfying chosen parameters: 214280

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

```

Command line parameters:
-MODEL=framed_n2p.model -DEV=xlp
-O=-cgnr1/vs/spot1/0509835992/r/unat.14012003.161528.4118/app_query.fasta.1.3932
-D=Published.Applications_AA -OEMT=faetan -SUFFIX=olig_rapid -MINMATCH=0.1
-LOOPC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATH=oligo
-TRANS=humanad cdt -LITER=45 -DOCALC=200 -THR.SCORE=quality -THR.MIN=1
-ALIGN=5 -MODE=LOCAL -OUTPMT=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=0509835992 -CGC=1_1_22 -ENGIN=14012003.161528.4118
-NCPU=6 -ICPU=3 -NO.ALXPY -NO.MMAP -LARGESUBSTR -ENG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -MAIN.TIMEOUT=30 -THREADS=1 -XGAOP=60 -XGAPEXT=60 -FGAOP=6
-DELOP=7 -YGAOP=60 -YGAPEXT=60 -DELOP=7

```

Database : Published_Applications_AA:*

- | | |
|-----|---|
| 1: | /cgn2.6/plodata/2/pubppa/US08_NEW_PUB_dep.* |
| 2: | /cgn2.6/plodata/2/pubppa/US08_NEW_PUB_dep.* |
| 3: | /cgn2.6/plodata/2/pubppa/US06_NEW_PUB_dep.* |
| 4: | /cgn2.6/plodata/2/pubppa/US06_NEW_PUB_dep.* |
| 5: | /cgn2.6/plodata/2/pubppa/US07_NEW_PUB_dep.* |
| 6: | /cgn2.6/plodata/2/pubppa/US07_PUBCOMB_dep.* |
| 7: | /cgn2.6/plodata/2/pubppa/US08_PUBCOMB_dep.* |
| 8: | /cgn2.6/plodata/2/pubppa/US08_NEW_PUB_dep.* |
| 9: | /cgn2.6/plodata/2/pubppa/US09_NEW_PUB_dep.* |
| 10: | /cgn2.6/plodata/2/pubppa/US09_PUBCOMB_dep.* |
| 11: | /cgn2.6/plodata/2/pubppa/US10_NEW_PUB_dep.* |
| 12: | /cgn2.6/plodata/2/pubppa/US10_PUBCOMB_dep.* |
| 13: | /cgn2.6/plodata/2/pubppa/US06_NEW_PUB_dep.* |
| 14: | /cgn2.6/plodata/2/pubppa/US06_PUBCOMB_dep.* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	8	2.5	143	12	US-10-062-254-568	Sequence 368, App
2	8	2.5	359	9	US-10-012-052-4	Sequence 4, App1
3	8	2.5	397	10	US-09-815-242-12461	Sequence 12461, A
4	7	2.2	42	10	US-09-864-761-41400	Sequence 41400, A

ALIGNMENTS

RESULT 1
US-10-062-254-368

; GENERAL INFORMATION:

: APPLICANT: Canoon, Edgar B
 : APPLICANT: Cahoon, Rebecca E
 : APPLICANT: Falco, Saverio Carl

APPLICANT: Hantke, Sabine S.

APPLICANT: L1, Zhongsen

APPLICANT: Morgante, Michele

APPLICANT: Odell, Joan

APPLICANT: Sakai, Hajime

APPLICANT: Zhu, Qun

```

; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/062,254
; CURRENT STRING DATE: 2009.09.01

```

PRIOR APPLICATION NUMBER: 09/630,346

PRIOR FILING DATE: 1999-07-30

C	5	7	2.2	52	10	US-09-864-761-335057	Sequence 35057,
C	6	7	2.3	54	10	US-09-864-761-46413	Sequence 46413,
C	7	7	2.3	60	10	US-09-864-761-45360	Sequence 45360,
C	8	7	2.3	64	9	US-09-796-692-1638	Sequence 1638, Ap
C	9	7	2.3	64	9	US-09-796-692-1978	Sequence 1978, Ap
C	10	7	2.2	71	12	US-10-078-929-1174	Sequence 1174, Ap
C	11	7	2.2	74	10	US-09-864-761-44212	Sequence 42412,
C	12	7	2.2	74	12	US-10-078-929-1176	Sequence 1176, Ap
C	13	7	2.3	87	9	US-09-796-692-1515	Sequence 1515, Ap
C	14	7	2.3	87	9	US-09-796-692-1610	Sequence 1610, Ap
C	15	7	2.3	87	9	US-09-796-692-1745	Sequence 1745, Ap
C	16	7	2.3	87	9	US-09-796-692-1758	Sequence 1758, Ap
C	17	7	2.3	87	9	US-09-796-692-1942	Sequence 1942, Ap
C	18	7	2.3	87	9	US-09-796-692-1964	Sequence 1964, Ap
C	19	7	2.3	87	9	US-09-796-692-2037	Sequence 2027, Ap
C	20	7	2.3	87	9	US-09-796-692-2000	Sequence 2040, Ap
C	21	7	2.3	132	10	US-09-811-731-10	Sequence 10, App
C	22	7	2.3	139	10	US-09-864-761-43779	Sequence 43779,
C	23	7	2.2	159	9	US-09-764-868-960	Sequence 960, App
C	24	7	2.2	159	10	US-09-780-017-2	Sequence 2, Appl
C	25	7	2.2	173	10	US-09-925-297-668	Sequence 668, Ap
C	26	7	2.2	239	9	US-09-738-626-3774	Sequence 3774, Ap
C	27	7	2.2	260	10	US-09-811-731-16	Sequence 16, App
C	28	7	2.3	260	12	US-10-000-954-10	Sequence 10, App
C	29	7	2.3	274	10	US-09-939-980-459	Sequence 499, App
C	30	7	2.3	294	10	US-09-815-242-5886	Sequence 5886, A
C	31	7	2.3	298	10	US-09-748-033-1	Sequence 1, Appl
C	32	7	2.3	299	10	US-09-815-242-13904	Sequence 13904,
C	33	7	2.3	299	10	US-09-815-242-11086	Sequence 11086,
C	34	7	2.2	358	10	US-09-815-242-1548	Sequence 5248, A
C	35	7	2.2	358	10	US-09-815-242-11363	Sequence 12163,
C	36	7	2.2	426	10	US-09-765-272-48	Sequence 48, App
C	37	7	2.2	458	10	US-09-954-314-2	Sequence 2, Appl
C	38	7	2.3	471	10	US-09-801-196-21	Sequence 21, App
C	39	7	2.3	471	10	US-09-801-196-32	Sequence 32, App
C	40	7	2.3	482	12	US-10-078-929-180	Sequence 180, App
C	41	7	2.3	545	10	US-09-908-988B-4	Sequence 4, Appl
C	42	7	2.3	548	10	US-09-915-806-3	Sequence 3, Appl
C	43	7	2.3	604	10	US-09-925-301-1026	Sequence 1026, A
C	44	7	2.2	610	9	US-09-992-998-113	Sequence 113, App
C	45	7	2.2	610	9	US-09-989-293A-113	Sequence 113, App

PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287
PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO 368
LENGTH: 143
TYPE: PRT
ORGANISM: Homo sapiens
US-10-062-254-368

Alignment Scores:
Pred. No.: 4.25 Length: 143
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 12

US-09-835-992A-21 (1-994) x US-10-062-254-368 (1-143)

OY 760 TTCCCGAAGCGCGCATTCCTT 783
|||||
DB 3 PheptoglialalaserSerphe 10

RESULT 2
US-10-012-055-4
Sequence 4, Application US/10012055
Patent No. US20020164750A1
GENERAL INFORMATION:

APPLICANT: Meyers, Rachel A.
APPLICANT: Rudolph-Owen, Laura A.
TITLE OF INVENTION: 56634, A NOVEL HUMAN PHOSPHATIDYLINOSITOL 4-PHOSPHATE 5-KINASE
FILE REFERENCE: 10448-114001
CURRENT APPLICATION NUMBER: US/10/012.055
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: 60/248,325
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 359
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus sequence
US-10-012-055-4

Alignment Scores:
Pred. No.: 3.92 Length: 359
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 9

US-09-835-992A-21 (1-994) x US-10-012-055-4 (1-359)

OY 337 TTGAAGAAAAAGCATTTTAAAA 360

|||||
DB 152 LeuTylsYlSAlaLeuLeuLys 159

RESULT 3
US-09-815-242-12461
Sequence 12461, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12461
LENGTH: 397
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12461

Alignment Scores:
Pred. No.: 3.89 Length: 397
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.55% Indels: 0
Gaps: 0
DB: 10

US-09-835-992A-21 (1-994) x US-09-815-242-12461 (1-397)

OY 597 GCACACCCACTAGCGCTACAGT 620
|||||
DB 350 GlyHisProLeuGlyAlaThrGly 357

RESULT 4
US-09-864-761-41400
Sequence 41400, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04

US-09-835-992a-21 (1-994) x US-09-864-761-35057 (1-52)

QY 455 TTGCTTTTACCACGAAGT 475
|||||
Db 36 LeuLeuPheTyrGlnArgThr 42

RESULT 6
US-09-864-761-46413

; Sequence 46413, Application US/09864761
; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aecmics-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 46413

; LENGTH: 54

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC020897.3

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3

; OTHER INFORMATION: SWISSPROT HIT: O32869, EVALU 7.00e-01

; OTHER INFORMATION: EST_HUMAN HIT: BF436304.1, EVALU 3.70e+00

; US-09-864-761-46413

Alignment Scores:

Pred. No.: 51.9

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.26% Indels: 0

DB: 10 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-864-761-46413 (1-54)

QY 782 AAGGAACGGCGCTTCGGG 762

|||||
Db 12 LysGluLeuAlaAlaSerGly 18

RESULT 7

US-09-864-761-45360

; Sequence 45360, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

; FILE REFERENCE: Aecmics-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 45360

; LENGTH: 60

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC013626.3

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EST_HUMAN HIT: F06450.1, EVALUATE 4.00e-26
OTHER INFORMATION: SWISSPROT HIT: P36607, EVALUATE 1.00e-08
US-09-864-761-45360

Alignment Scores:

Pred. No.: 51.4 Length: 60
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-864-761-45360 (1-60)

OY 657 CTTCAGCTGCAGCAGCTT 637

DB 49 LeuSerAlaAlaSerMgVal 55

RESULT 8

US-09-796-692-1638
Sequence 1638, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1638
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(64)
OTHER INFORMATION: Xaa - Any amino acid
US-09-796-692-1638

Alignment Scores:

Pred. No.: 51.1 Length: 64
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-796-692-1638 (1-64)

OY 663 CTTCCTCCTGCAGTGCAGC 643

DB 24 LeuProLeuSerAlaAlaSer 30

RESULT 9

US-09-796-692-1978
Sequence 1978, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Galger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1978
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(64)
OTHER INFORMATION: Xaa - Any amino acid
US-09-796-692-1978

Alignment Scores:

Pred. No.: 51.1 Length: 64
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-796-692-1978 (1-64)

OY 663 CTTCCTCCTGCAGTGCAGC 643

DB 24 LeuProLeuSerAlaAlaSer 30

RESULT 10

US-10-078-929-174
; Sequence 174, Application US/10078929
; Patent No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Ramodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; FILE REFERENCE: Stress Response
; FILE REFERENCE: BBI357 US NA
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US/10/078,929
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 174
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
US-10-078-929-174
Alignment Scores:
Pred. No.: 50.7 Length: 71
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 12 Gaps: 0
US-09-835-992a-21 (1-994) x US-10-078-929-174 (1-71)
QY 780 CTTTGAAGTCATCAAAATT 800
Db 43 LeuLeuGluLeuLeuLeuLeuLeu 49
RESULT 11
US-09-864-761-42412
; Sequence 42412, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42412
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 297197.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EST_HUMAN HIT: AA099217.1, EVALUATE 7.00e-19
; OTHER INFORMATION: SWISSPROT HIT: Q92793, EVALUATE 6.40e+00
US-09-864-761-42412
Alignment Scores:
Pred. No.: 50.5 Length: 74
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 10 Gaps: 0
US-09-835-992a-21 (1-994) x US-09-864-761-42412 (1-74)
QY 809 TTCCAAACCAAGCTCGCAAGT 829
Db 67 PheGlnProSerSerAlaSer 73
RESULT 12
US-10-078-929-176
; Sequence 176, Application US/10078929
; Patent No. US20020152497A1


```

Alignment Scores:
Pred. No.: 50.5 Length: 74
Percent: 7.00 Matches: 7
Score Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 12 Gaps: 0

US-09-835-992A-21 (1-994) x US-10-078-929-176 (1-74)
OY 780 CTTTGAACATCAATCAAAATT 800
Db 45 LeuLengIuLeuIleLySile 51

RESULT 13
US-09-796-692-1515
; Sequence 1515, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545

```

Alignment Scores:	
Pred. No.:	49.8
Score:	7.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.26%
DB:	9
Gaps:	0
Length:	8
Matches:	7
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

```

US-09-835-992a-21 (1-994) x US-09-796-692-1515 (1-87)
OY      663  CTTCCCTTCAGCTGCACG 643
          |||||||
Db      47  LeuProLeuSerAlaAlaSer 53

RESULT 14
US-09-796-692-1610
: Sequence 1610, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Gaiger, Alexander
: APPLICANT: Algate, Paul A.
: ATTORNEY: Mannion, Jane
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
: FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
: FILE REFERENCE: 2077.001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: 60/222,903

```

;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1610
;; LENGTH: 87
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-796-692-1610

Alignment Scores:
Pred. No.: 49.8 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-796-692-1610 (1-87)

OY 663 CTTCCTCTCAGTCGCAGC 643

DB 47 LeuProLeuSerAlaAlaLaser 53

RESULT 15
US-09-796-692-1745

;; Sequence 1745, Application US/09796692
;; Publication No. US20020198362A1

;; GENERAL INFORMATION:

;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.

;; APPLICANT: Mannion, Jane

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES

;; FILE REFERENCE: 2077.001200
;; CURRENT APPLICATION NUMBER: US/09/796,692

;; PRIOR FILING DATE: 2001-03-01

;; PRIOR APPLICATION NUMBER: 60/186,126

;; PRIOR FILING DATE: 2000-03-01

;; PRIOR APPLICATION NUMBER: 60/190,479

;; PRIOR FILING DATE: 2000-03-17

;; PRIOR APPLICATION NUMBER: 60/200,545

;; PRIOR FILING DATE: 2000-04-27

;; PRIOR APPLICATION NUMBER: 60/200,303

;; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: 60/200,779

;; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: 60/200,999

;; PRIOR FILING DATE: 2000-05-01

;; PRIOR APPLICATION NUMBER: 60/202,084

;; PRIOR FILING DATE: 2000-05-04

;; PRIOR APPLICATION NUMBER: 60/206,201

;; PRIOR FILING DATE: 2000-05-22

;; PRIOR APPLICATION NUMBER: 60/218,950

;; PRIOR FILING DATE: 2000-07-14

;; PRIOR APPLICATION NUMBER: 60/222,903

;; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: 60/223,416

;; PRIOR FILING DATE: 2000-08-04

;; PRIOR APPLICATION NUMBER: 60/223,378

;; PRIOR FILING DATE: 2000-08-07

;; NUMBER OF SEQ ID NOS: 9597

;; SOFTWARE: FastSeq for Windows Version 3.0

;; SEQ ID NO 1745

;; LENGTH: 87

;; TYPE: PRT

;; ORGANISM: Homo sapiens

US-09-796-692-1745

Alignment Scores:

Pred. No.: 49.8 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-21 (1-994) x US-09-796-692-1745 (1-87)

OY 663 CTTCCTCTCAGTCGCAGC 643

DB 47 LeuProLeuSerAlaAlaLaser 53

Search completed: January 14, 2003, 18:08:20
Job time : 11.5164 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:45 : Search time 33.3073 Seconds
(without alignments)
5737.940 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 314
Sequence: 1 cccaccagctgctcctcag.....ccaccctgggctgggagat 994

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 56536

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p_model -DEV=x1p
-Q/cgcn2.1/USFTO.spool/US09835992/rnatat_14012003_161526_4038/app_query.fasta.1.3932
-DB=PIR_73 -QPM=fastan -SUFFIX=olg.rpr -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LST=45
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992 @cgn.1.1.192 @rnatat.14012003.161526.4038 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOC -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FAPOP=6 -FAPEXT=7
-YAPOP=60 -YAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	223	71.0	547	2	I38205	sterol carrier pro
2	84	26.8	546	2	B40407	sterol carrier pro
3	50	15.9	547	2	JU0157	sterol carrier pro
4	49	15.6	547	2	A39368	sterol carrier pro
5	33	10.5	547	2	S34744	sterol carrier pro
6	24	7.6	401	2	E70884	probable lipid-tra
7	22	7.0	412	2	T27202	hypothetical prote
8	10	3.2	387	2	B90436	hypothetical prote
9	9	2.9	246	2	T45470	ATP-binding ABC tr
c 10	9	2.9	351	2	A96559	hypothetical prote
11	9	2.9	383	2	G69205	lipid-transfer pro
12	8	2.5	143	2	A40015	sterol carrier pro
13	8	2.5	231	2	T19267	hypothetical prote
14	8	2.5	267	2	T35889	hypothetical prote

15	8	2.5	308	2	S58137	gene 7 protein - p
c 16	8	2.6	314	2	T05993	probable peroxidase
c 17	8	2.6	375	1	S66272	alcohol dehydrogen
18	8	2.5	379	2	C83278	probable acyl-CoA
19	8	2.5	383	2	E84249	3-ketoacyl-CoA thl
20	8	2.5	384	2	B41838	two-component sens
21	8	2.5	386	2	B70604	probable fadA6 pro
22	8	2.5	387	2	F69304	3-ketoacyl-CoA thl
23	8	2.5	388	2	H69370	3-ketoacyl-CoA thl
24	8	2.5	391	2	B69252	3-ketoacyl-CoA thl
25	8	2.5	394	2	B89786	hypothetical prote
26	8	2.5	395	2	H90421	hypothetical prote
27	8	2.5	398	1	B64383	molybdenum cofacto
28	8	2.5	400	2	A87427	thiolase family pr
29	8	2.5	401	2	E83427	probable acyl-CoA
30	8	2.5	401	2	D87258	fatty oxidation co
31	8	2.5	402	2	B97420	probable acyl-CoA
32	8	2.5	402	2	AC2638	acyl-CoA thiolase
33	8	2.5	403	2	C70815	probable beta-keto
34	8	2.5	403	2	E87179	probable beta-keto
35	8	2.5	404	2	T35428	probable acetyl co
36	8	2.5	404	2	T35256	probable thiolase
37	8	2.5	407	2	E87634	thiolase family pr
38	8	2.5	412	2	B70582	probable lipid tra
39	8	2.5	414	2	S36838	acetyl-CoA C-acylt
40	8	2.5	414	2	T52165	acetyl-CoA C-acylt
41	8	2.5	417	2	AG3571	acetyl-CoA C-acylt
42	8	2.5	426	2	T01790	protoporphyrin IX
43	8	2.5	430	2	S57792	acetyl-CoA C-acylt
44	8	2.5	436	2	D84782	probable prolone t
45	8	2.6	440	2	T24837	hypothetical prote

ALIGNMENTS

RESULT 1
I38205
sterol carrier protein X/sterol carrier protein-2 - human
C:Species: Homo sapiens (man)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 08-Sep-2000
C:Accession: I38205
R:Onda, T.; Kennert, H.; Pfeiffer, S.M.; He, Z.; Yamamoto, R.; Holt, J.A.; Billheimer, Genomics 24, 370-374, 1994
A:Title: The structure of the human sterol carrier protein X/sterol carrier protein 2
A:Reference number: I38205; MUID:95213031; PMID:7698762
A:Accession: I38205
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-547 <RES>
A:Cross-references: EMBL:U11313; NID:9532077; PID:9532079
A:Note: nucleotide sequence is not complete
C:Genetics:
A:Gene: GDB:SCP2
A:Cross-references: GDB:126877; OMIM:184755
A:Map position: IP32-1p32
A:Introns: 23/3; 43/1; 67/1; 111/1; 132/3; 175/1; 196/2; 225/2; 275/3; 325/1; 361/1;

Alignment Scores:

Pred. No.: 5.78e-224 Length: 547
Score: 223.00 Matches: 223
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 71.02% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-21 (1-994) x I38205 (1-547)

QY 3 CACCAAGTGTCTCCTCAGATGTTGGATGATCGAAGAAACATATGAAATATGCA 62
Db 156 HlspProvaAlapProGlnMetPheGlyTyrAlaGlySGlnHlsmecGlnulstYrftly 175
QY 63 ACNAAATGTGACCTTTGCAAAATTTGGATGAGAAATTCATTAACATTCAGTTAATAC 122

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.92% Indels: 0
 Gaps: 0

US-09-835-992a-21 (1-994) x JU0157 (1-547)

OY 522 ACCTGCTGATAGACAGATATATCATATGAGAGAAAGTGGCTCATTAATCTAGTGT 581
 |||||||
 Db 329 ThrLeuValAspArgGlyAspAsnThrTyrGlyGlyLysTrpValIleAsnProSerGly 348
 |||||||

OY 582 GGAATGATTTCAAGGACACCCCTAGCGGCTACAGTCTTGTCTGCTGACGACACTC 641
 |||||||
 Db 349 GlyLeuLeuSerGlyGlyHisProLeuGlyValThrGlyLeuAlaGlnCysAlaGluLeu 368
 |||||||

OY 642 TGCTGACGCTGAGAGGGAACCGGAAAA 671
 |||||||
 Db 369 CysTrpPoleuArgGlyGlyValAlaGlyLys 378
 |||||||

RESULT 4

A39368
 A:Species: Rat:us norvegicus (Norway rat)
 C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 01-Dec-2000
 C:Accession: A39368; B39054; A29366; A27661; S17842; A34584; A34635
 R:Morl, T.; Tsukamoto, T.; Mori, H.; Tashiro, Y.; Fujiki, Y.
 Proc. Natl. Acad. Sci. U.S.A. 88, 4338-4342, 1991

A:Title: Molecular cloning and deduced amino acid sequence of nonspecific lipid transfer sequence of nonspecific lipid transfer protein as its C-terminal part.
 A:Reference number: A39368; MUID:91239563; PMID:2034675

A:Accession: A39368
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-547 <MOR>

A:Cross-references: GB:62763; NID:g202552; PIDN:AAA0622.1; PID:g202553
 R:Seedorf, U.; Asmann, G.
 J. Biol. Chem. 266, 630-636, 1991

A:Title: Cloning, expression, and nucleotide sequence of rat liver sterol carrier protein
 A:Reference number: A39054; MUID:91093192; PMID:1985920

A:Accession: A39054
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 405-547 <SEF>

A:Cross-references: GB:J05716
 A:Accession: B39054
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-11, 'P', 13-49, 'R', 51-264, 'V', 266-547 <SE2>

A:Cross-references: GB:J05717
 R:Paluszyn, A.; Noland, B.J.; Bazan, J.F.; Pletterick, R.J.; Scallan, T.J.
 J. Biol. Chem. 262, 13219-13227, 1987

A:Title: Primary sequence and structural analysis of sterol carrier protein 2 from rat
 A:Reference number: A29366; MUID:88007528; PMID:3115977

A:Accession: A29366
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 425-546 <PAS>

R:Morris, H.R.; Larsen, B.S.; Billheimer, J.T.
 Biochem. Biophys. Res. Commun. 154, 476-482, 1988

A:Title: A mass spectrometric study of the structure of the structure of sterol carrier protein SCP-2 from
 A:Reference number: A27661; MUID:88280812; PMID:3395344

A:Accession: A27661
 A:Molecule type: protein
 A:Residues: 425-426, 'SV', 429-435, 'V', 437-445, 'D', 447-449, 'Q', 451-487, 'N', 489-515, 'T', 517
 R:Ossendorp, B.C.; van Heusden, G.P.H.; de Beer, A.L.J.; Bos, K.; Schouten, G.L.; Wirtz, Eur. J. Biochem. 201, 233-239, 1991

A:Title: Identification of the cDNA clone which encodes the 58-kDa protein containing the
 th rat peroxisomal and mitochondrial 3-oxoacyl-CoA thiolases.
 A:Reference number: S17842; MUID:92007881; PMID:1915365

A:Accession: S17842
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 2-49, 'R', 51-547 <OSS>

A:Cross-references: EMBL:X60654; NID:956771; PIDN:CAA43061.1; PID:g56773

R:Billheimer, J.T.; Strehl, L.L.; Davis, G.L.; Strauss III, J.F.; Davis, L.G.
 DNA Cell Biol. 9, 159-165, 1990

A:Title: Characterization of a cDNA encoding rat sterol carrier protein-2.
 A:Reference number: A34584; MUID:90253610; PMID:2340090

A:Accession: A34584
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 275-547 <BIL>

A:Cross-references: GB:M34728; NID:g206871; PIDN:AAA42120.1; PID:g206872
 R:Ossendorp, B.C.; van Heusden, G.P.H.; Wirtz, K.W.A.
 Biochem. Biophys. Res. Commun. 168, 631-636, 1990

A:Title: The amino acid sequence of rat liver non-specific lipid transfer protein (st
 A:Reference number: A34635; MUID:90241231; PMID:2334427

A:Accession: A34635
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 33-49, 'R', 51-547 <OS2>

A:Cross-references: GB:M58287; NID:g205769; PIDN:AAA41726.1; PID:g205770
 C:Keywords: alternative splicing; mitochondrion

Alignment Scores:

Pred. No.: 2,97e-42 Length: 547
 Score: 49.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.61% Indels: 0
 Gaps: 0

US-09-835-992a-21 (1-994) x A39368 (1-547)

OY 525 CTGCTGATAGACAGATATATCATATGAGAGAAAGTGGCTCATTAATCTAGTGTGCA 584
 |||||||
 Db 330 LeuValAspArgGlyAspAsnThrTyrGlyGlyLysTrpValIleAsnProSerGlyGly 349
 |||||||

OY 585 CTGATTCGAAGGACACCCCTAGCGGCTACAGTCTTGTCTGCTGACGACACTGCG 644
 |||||||
 Db 350 LeuLeuSerGlyGlyHisProLeuGlyValThrGlyLeuAlaGlnCysAlaGluLeuCys 369
 |||||||

OY 645 TGCAAGCTGAGAGGGAACCGGAAAA 671
 |||||||
 Db 370 TrpGlnLeuArgGlyGlyValAlaGlyLys 378
 |||||||

RESULT 5

S34744
 A:Species: Gallus gallus (chicken)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000
 C:Accession: S34744
 R:Peifer, S.M.; Sakurai, N.; Ryan, A.; Johnson, A.L.; Deeley, R.G.; Billheimer, J.T.
 Arch. Biochem. Biophys. 304, 287-293, 1993

A:Title: Chicken sterol carrier protein 2/sterol carrier protein x: cDNA cloning reve
 A:Reference number: S34744; MUID:93312016; PMID:8323294

A:Accession: S34744
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-547 <PEF>

A:Cross-references: GB:I09231; NID:g304422; PIDN:AAA02488.1; PID:g304423

Alignment Scores:
 Pred. No.: 1.52e-25 Length: 547
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.51% Indels: 0
 Gaps: 0

US-09-835-992a-21 (1-994) x S34744 (1-547)

OY 531 GATAGAGAGATATATCATATGAGAGAAAGTGGCTCATTAATCTAGTGTGACTGATT 590
 |||||||
 Db 335 AspArgGlyAspAsnThrTyrGlyGlyLysTrpValIleAsnProSerGlyGlyLeuLe 354
 |||||||

OY 591 TCAGAGGACACCCCTAGCGGCTACAGTCTTGTCTGACG 629
 |||||||

DB 355 SerLySGlyHisProLeuGlyAlaThrGlyLeuAlaGln 367
|||||
RESULT 6
E70884
probable lipid-transfer protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: E70884
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: E70884
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-401 <COL>
A: Cross-references: GB:AL008967; GB:AL123456; NID: g3261491; PIDN: CA15585.1; PID: g262431
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: ltp1
C: Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:
Pred. No.: 3,97e-16 Length: 401
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.64% Indels: 0
DB: Gaps: 0

US-09-835-992A-21 (1-994) x E70884 (1-401)

QY 570 AATCCAGTGGTGGATTCCTCAAGGACACCCATGAGGCTACAGTCTGTCTGAG 629
|||||
DB 338 AsnProSerGlyGlyLeuIleSerLySGlyHisProLeuGlyAlaThrGlyLeuAlaGln 357
|||||
QY 630 TGTGCGAGAACTC 641
|||||
DB 358 CysAlaGlyLeu 361
|||||
RESULT 7
T27202
hypothetical protein Y57A10C.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T27202
R: Smye, R.
submitted to the EMBL Data Library, June 1998
A: Reference number: Z20327
A: Accession: T27202
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-412 <WIL>
A: Cross-references: EMBL: AL023847; PIDN: CA19548.1; GSPDB: GN00020; CESP: Y57A10C.6
A: Experimental source: clone Y57A10C
C: Genetics:
A: Gene: CESP: Y57A10C.6
A: Map position: 2
A: Introns: 62/1; 104/1; 171/1
C: Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:
Pred. No.: 4,86e-14 Length: 412
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.01% Indels: 0
DB: Gaps: 0

US-09-835-992A-21 (1-994) x T27202 (1-412)

QY 540 GATATATCATATGAGAGAAAGTGGTCATTAATCTAGTGGACCTGATTCAAGCGA 599
|||||
DB 332 AspAsnThrGlyGlyGlySerPvalIleAsnProSerGlyLeuIleSerLySGly 351
|||||
QY 600 CACCCA 605
|||||
DB 352 HisPro 353
|||||
RESULT 8
B90436
hypothetical protein acab-7 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: B90436
R: She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A: Description: Sulfolobus solfataricus complete genome.
A: Reference number: A99139
A: Accession: B90436
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-387 <KUR>
A: Cross-references: GB:AE006641; NID: g13815933; PIDN: AAK42745.1; GSPDB: GN00155
C: Genetics:
A: Gene: acab-7
C: Superfamily: probable 3-ketoacyl-CoA thiolase MTH1177

Alignment Scores:
Pred. No.: 0.167 Length: 387
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.18% Indels: 0
DB: Gaps: 0

US-09-835-992A-21 (1-994) x B90436 (1-387)

QY 594 AAGGACACCCACTGAGCGCTACAGTCTT 623
|||||
DB 333 LysGlyHisProLeuGlyAlaThrGlyLeu 342
|||||
RESULT 9
T45470
ATP-binding ABC transport protein [imported] - Erysipelothrix rhusiopathiae
C:Species: Erysipelothrix rhusiopathiae
C>Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 17-Mar-2000
C:Accession: T45470
R: Antorj, R.P.; Turner, B.; Eamens, G.J.; Delaney, S.F.; Chin, J.C.
submitted to the EMBL Data Library, March 1996
A: Description: Cloning and nucleotide sequencing of ew1A, the gene encoding an immuno
A: Reference number: Z22979
A: Accession: T45470
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-246 <ANT>
A: Cross-references: EMBL: U52850; PIDN: AA801157.1
A: Experimental source: strain VAS 229
C: Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Alignment Scores:
Pred. No.: 1.97 Length: 246
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.90% Indels: 0
DB: Gaps: 0

US-09-835-992A-21 (1-994) x T45470 (1-246)

OY 782 AACGAACGCGGCTTCCGCAAAACC 756
 DB 184 LygGlutLeuAlaAlaSerGlyLysTr 192

RESULT 10

A96559

hypothetical protein P5F19.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A96559

R:Ethelberg, A.; Ecker, J.R.; Palm, C.J.; Federle, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,

Rizzo, T.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A06141; MUID:21016719; PMID:11130712

A:Accession: A96559

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <STO>

A:Cross-references: GB:AE005173; NID:94220443; PIDN:AD12670.1; GSPDB:GN00141

C:Genetics:

A:Gene: P5F19.1

A:Map position: 1

Alignment Scores:

Pred. No.:	1.87	Length:	351
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.87%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-21 (1-994) x A96559 (1-351)

OY 150 AGTTAGTGAAGTGCATCTTAA 176

DB 320 SerLunspLcValMetAlaSerLys 328

RESULT 11

G69205

lipid-transfer protein (sterol or nonspecific) - Methanobacterium thermoautotrophicum (S

C:Species: Methanobacterium thermoautotrophicum

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999

C:Accession: G69205

R:Smith, D.R.; Doucette-Stamm, L.A.; Delouphery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spodoforo, R.; Vicaltre, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: G69205

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1383 <MTH>

A:Cross-references: GB:AE000857; GB:AE000666; NID:92621876; PIDN:AAB85293.1; PID:9262188

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH793

C:Superfamily: probable 3-ketolacyl-CoA thiolase MTH1177

Alignment Scores:

Pred. No.:	1.85	Length:	383
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.87%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-21 (1-994) x G69205 (1-383)

OY 645 TGCCAGCTGAGAGGAGCCGGA 671

DB 347 TrpGlnLeuArgGlyGlnAlaGlyLys 355

RESULT 12

A40015

steroid carrier protein 2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 02-Sep-2000

C:Accession: A40015

R:Monocchi, D.; Pastuszyn, A.; Scallan, T.J.

J. Biol. Chem. 266, 9885-9892, 1991

A:Title: cDNA sequence and bacterial expression of mouse liver steroid carrier protein

A:Reference number: A40015; MUID:91236770; PMID:1709640

A:Accession: A40015

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-143 <MON>

A:Cross-references: GB:M62361; NID:9200941; PIDN:AAA4009.1; PID:9200942

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK892.2

C:Keywords: mitochondrion

Alignment Scores:

Pred. No.:	23.6	Length:	143
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-21 (1-994) x A40015 (1-143)

OY 760 TTCGCCGAGCCGCGAGTTCCTT 783

DB 3 PheProGlnAlaAlaSerSerPhe 10

RESULT 13

T19267

hypothetical protein C14B1.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19267

R:Harris, B.

submitted to the EMBL Data Library, September 1994

A:Reference number: Z19099

A:Accession: T19267

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-231 <NTL>

A:Cross-references: EMBL:Z27139; PIDN:CAA85488.1; GSPDB:GN00021; CESP:C14B1.8

A:Experimental source: clone C14B1

C:Genetics:

A:Gene: CESP:C14B1.8

A:Map position: 3

A:Introns: 25/1; 58/1; 117/2; 146/2; 181/2

Alignment Scores:

Pred. No.:	22	Length:	231
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-21 (1-994) x T19267 (1-231)

OY 331 GCTGCTTGAAGAAAGCATTA 354

DB 143 AlaArgLeuLysLysLysAlaLeu 150

RESULT 14

T35889
 hypothetical protein SC9B10.20c - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C:Accession: T35889
 R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 Submitted to the EMBL Data Library, November 1997
 A:Reference number: Z21592
 A:Accession: T35889
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-267 <OLI>
 A:Cross-references: EMBL:AL009204; PIDN:CA15810.1; GSPDB:GN00070; SCOEDB:SC9B10.20c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC9B10.20c
 C:Superfamily: Mycobacterium hypothetical protein RV0911

Alignment Scores:

Pred. No.:	21.6	Length:	267
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992a-21 (1-994) x T35889 (1-267)

OY 468 AACGACCTCTTACTTATGACGA 491
 |||||

Db 154 ASngtuleuLeuThrTyrGluAla 161

RESULT 15

S58137

gene 7 protein - phage SPPI

C:Species: phage SPPI

C:Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 11-May-2000

C:Accession: S58137; S24456; T42271

R:Becker, B.; Gassel, W.; Tavares, P.; Lurz, R.; Alonso, J.C.

submitted to the EMBL Data Library, July 1995

A:Description: Head morphogenesis of the Bacillus subtilis bacteriophage SPPI.

A:Reference number: S58137

A:Accession: S58137

A:Molecule type: DNA

A:Residues: 1-308 <BEC>

A:Cross-references: EMBL:X89721; NID:g1052805; PIDN:CA61865.1; PID:g1052806

R:Chai, S.; Bravo, A.; Lueder, G.; Nedlin, A.; Trautner, T.A.; Alonso, J.C.

J. Mol. Biol. 224, 87-102, 1992

A:Title: Molecular analysis of the Bacillus subtilis bacteriophage SPPI region encompassing the head morphogenesis genes

A:Reference number: S24450; MUID:92194332; PMID:1548711

A:Accession: S24456

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-43 <CHA>

A:Cross-references: EMBL:X56064; NID:g15464; PIDN:CAA39542.1; PID:g15471

R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.

Gene 204, 201-212, 1997

A:Title: The complete nucleotide sequence and functional organization of Bacillus subtilis bacteriophage SPPI

A:Reference number: Z22137; MUID:98094274; PMID:9434185

A:Accession: T42271

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-308 <ALO>

A:Cross-references: EMBL:X97918; PIDN:CA66581.1

Alignment Scores:

Pred. No.:	21.1	Length:	308
Score:	8.00 <td>Matches:</td> <td>8</td>	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.55%	Indels:	0

DB: 2 Gaps: 0

US-09-835-992a-21 (1-994) x S58137 (1-308)

OY 85 AATTGGATGGAAAAATCATTAAC 108
 |||||

Db 230 LysleuaspGlyLysleileAsn 237

Search completed: January 14, 2003, 17:40:45
 Job Time : 39.3073 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 15.349 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992a-21

Perfect score: 314
Sequence: 1 ctcaccagctgcctccag.....ccaccctgggtgttgat 994

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Deloxt 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+n2p.model -DEV-x1p
-O/cg92.1/USPTO.spool/US09835992/runeat_14012003_161525_4014/app_query.fasta_1.3932
-DB-SwissProt_40 -QWMT-fastan -SUFFIX-oligo -MIMATCH-0.1 -LOOPEXT-0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MARIK-oligo -TRANS-human40.cd1
-LIST-45 -DOCALLIN-200 -THR.SCORE-quality -THR.MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFM-pto -NOR-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USR-US09835992.ecgn_1.1.77-eunal.14012003_161525_4014 -NCPU-6 -ICPU-3
-NO.XLPHY -NO.MMP -LARGEQUERY -NEG.SCORES-0 -WAIT -LONGLOG -DEV.TIMEOUT-120
-MARN.TIMEOUT-30 -THREDS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAPEXT-60 -DELOP-6 -DELEXT-7

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	223	71.0	547 1	NLTP_HUMAN
2	50	15.9	547 1	NLTP_MOUSE
3	49	15.6	547 1	NLTP_RAT
4	33	10.5	547 1	NLTP_CHICK
5	8	2.5	151 1	YCX2_OENHO
6	8	2.6	374 1	ADH1_APTAU
7	8	2.5	384 1	VANS_ENTFC
8	8	2.5	398 1	Y666_METJA
9	8	2.5	414 1	THIK_YARLI
10	8	2.5	426 1	CHLI_TOBAC
11	8	2.6	455 1	POLG_MWY27
12	8	2.5	1005 1	P436_CHLTR
13	8	2.6	1016 1	POLG_MWY27
14	7	2.2	126 1	PAND_ECOLI
15	7	2.2	126 1	PAND_ECOLI
16	7	2.2	126 1	PAND_ECOLI
17	7	2.2	132 1	PAND_YERPE
18	7	2.2	132 1	PAND_YERPE
				MLE_TODPA

19	7	2.2	161 1	PIN_BPT4	P07068 bacterioph
20	7	2.2	167 1	YW20_MARPO	P38463 marchantia
21	7	2.3	169 1	PEP3_SCHPO	O10143 schizosacch
22	7	2.3	173 1	GRAM_MOUSE	P51437 mus musculu
23	7	2.2	199 1	HMG1_MOUSE	P40630 mus musculu
24	7	2.2	218 1	CD53_MOUSE	O61451 mus musculu
25	7	2.3	233 1	GP42_RAT	P23505 rattus norv
26	7	2.3	233 1	SN20_HUMAN	O60248 homo sapien
27	7	2.2	248 1	DSBG_ECO57	P58320 escherichia
28	7	2.2	248 1	DSBG_ECO57	P77202 escherichia
29	7	2.2	272 1	PRRF_DEIRA	O97963 delnoccocus
30	7	2.3	290 1	TYSY_HSVAT	P12462 herpesvirus
31	7	2.3	305 1	PEX2_MOUSE	P55098 mus musculu
32	7	2.3	305 1	PEX2_RAT	P24392 rattus norv
33	7	2.3	316 1	YK95_MYCVU	O10704 mycobacteri
34	7	2.2	322 1	YK46_AOUAE	O67500 aquilex aeo
35	7	2.2	330 1	OUTG_EMERT	O67500 aquilex aeo
36	7	2.2	334 1	SRB7_CAREL	P23416 emeritella
37	7	2.2	355 1	KARG_HONGA	P54142 caenorhabdi
38	7	2.2	356 1	KARG_SCHAM	P14208 homarus gam
39	7	2.2	367 1	LHX8_MOUSE	P91798 schistocerc
40	7	2.3	389 1	G68A_DROME	O35652 mus musculu
41	7	2.2	399 1	CHS2_HORVU	O97960 drosophila
42	7	2.2	421 1	PSMR_METAC	O96562 hordium vul
43	7	2.2	421 1	TIG_STRPN	O81188 methanosarc
44	7	2.2	449 1	MURF_RICPR	O97959 streptococc
45	7	2.2	450 1	MEI4_YEAST	O05953 rickettsia
					P29467 saccharomyc

ALIGNMENTS

RESULT 1

ID NLTP_HUMAN STANDARD: PRT: 547 AA.

AC P22307; Q16622; Q15432; Q99430; 01-AUG-1991 (Rel. 19, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Non-specific lipid-transfer protein, mitochondrial precursor (NSL-TP)

DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

DE (SCP).

Sequence from N.A.
MEDLINE-95213031; PubMed-7698762;
OBDA T., Rennett H., Pfeiffer S.M., He Z., Yamamoto R., Holt J.A.,
Billheimer J.T., Straus J.F. III;
"The structure of the human sterol carrier protein X/sterol carrier
protein 2 gene (SCP2)."
Genomics 24:370-374(1994).
[2]
SEQUENCE FROM N.A.
MEDLINE-92029618; PubMed-1718316;
He Z., Yamamoto R., Furtch E.E., Schantz L.J., Naylor S.L., George H.,
Billheimer J.T., Straus J.F. III;
"CDNAs encoding members of a family of proteins related to human
sterol carrier protein 2 and assignment of the gene to human
chromosome 1 p21-pter."
DNA Cell Biol. 10:559-569(1991).
[3]
SEQUENCE OF 405-547 FROM N.A.
TISSUE-Liver;
MEDLINE-9110550; PubMed-1703300;
Yamamoto R., Kallen C.B., Babalola G.O., Rennett H., Billheimer J.T.,
Straus J.F. III;
"Cloning and expression of a cDNA encoding human sterol carrier
protein 2.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:463-467(1991).
 RP [4]
 RN SEQUENCE OF 405-547 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93131254; PubMed=1483685;
 RA Yamamoto R.;
 RT "Localization of human sterol carrier protein 2 gene and CDNA
 expression in COS-7 cell.";
 RL Hokkaido Igaku Zasshi 67:839-848(1992).
 RN [5]
 RP STRUCTURE BY NMR OF SCP2.
 RX MEDLINE=94063072; PubMed=8243660;
 RA Szyerski T., Scheek S., Johansson J., Assmann G., Seedorf U.,
 Wuehrlich K.;
 RT "NMR determination of the secondary structure and the
 three-dimensional polypeptide backbone fold of the human sterol
 carrier protein 2.";
 RL FEBS Lett. 335:18-26(1993).
 CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
 PHOSPHOLIPIDS, CHOLESTEROL AND GANGLIOSIDES BETWEEN MEMBRANES. MAY
 PLAY A ROLE IN REGULATING STEROIDOGENESIS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
 WITH MITOCHONDRIA ESPECIALLY IN STEROIDOGENIC TISSUES. SCP-X
 RESIDES IN THE PEROXISOME.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, SCPx (shown here) and SCP2; are
 produced by alternative initiation.
 CC -1- TISSUE SPECIFICITY: LIVER, FIBROBLASTS, AND PLACENTA.
 CC -1- DISEASE: SCP2 IS PRESENT IN LOW LEVELS IN SUBJECTS WITH ZELLWEGER
 SYNDROME (CEREBRO-HEPATIC-RENAL SYNDROME), WHOSE CELLS ARE
 DEFICIENT IN PEROXISOMES AND WHO HAVE AN ASSOCIATED IMPAIRMENT IN
 PLASMA/OGN AND BILE ACID SYNTHESIS AND CATABOLISM OF PHYTANIC
 ACID AND VERY LONG-CHAIN FATTY ACIDS.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLASE
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U11313; AAB41286.1; -
 DR EMBL: U11297; AAB41286.1; JOINED.
 DR EMBL: U11299; AAB41286.1; JOINED.
 DR EMBL: U11300; AAB41286.1; JOINED.
 DR EMBL: U11301; AAB41286.1; JOINED.
 DR EMBL: U11302; AAB41286.1; JOINED.
 DR EMBL: U11303; AAB41286.1; JOINED.
 DR EMBL: U11304; AAB41286.1; JOINED.
 DR EMBL: U11305; AAB41286.1; JOINED.
 DR EMBL: U11306; AAB41286.1; JOINED.
 DR EMBL: U11307; AAB41286.1; JOINED.
 DR EMBL: U11308; AAB41286.1; JOINED.
 DR EMBL: U11309; AAB41286.1; JOINED.
 DR EMBL: U11310; AAB41286.1; JOINED.
 DR EMBL: U11311; AAB41286.1; JOINED.
 DR EMBL: U11312; AAB41286.1; JOINED.
 DR EMBL: M75883; AAA03557.1; -
 DR EMBL: M75884; AAA03558.1; ALT_INIT.
 DR EMBL: M55421; AAA03559.1; ALT_SEQ.
 DR EMBL: S52450; AAB24921.1; -
 DR HSSP: P27796; 1PXT.
 DR PIR: A39010; A39010.
 DR HSP: 184755; -
 DR GENE: HGNC:10606; SCP2.
 DR InterPro: IPR003033; SCP2.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF00108; thiolase; 1.
 DR Pfam: PF02036; SCP2; 1.
 DR Pfam: PF02803; thiolase; 1.
 DR PROSITE: PS00098; THIOLASE_1; 1.

DR PROSITE: PS00737; THIOLASE_2; 1.
 DR PROSITE: PS00099; THIOLASE_3; FALSE_NEG.
 DR PROSITE: PS00342; MICROBODIES_CTER; 1.
 KW Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
 KM Alternative initiation.
 FT CHAIN 1 547
 FT TRANSIT 405 424
 FT CHAIN 425 547
 FT INIT_MET 405 405
 FT ACT_SITE 94 94
 FT SITE 494 494
 FT SITE 545 547
 FT CONFLICT 10 10
 FT CONFLICT 393 393
 FT CONFLICT 472 472
 FT CONFLICT 482 482
 FT CONFLICT 501 501
 FT CONFLICT 522 522
 FT SEQUENCE 547 AA; 58993 MW; 29F7551465C7143A CRC64;
 60
 Alignment Scores:
 Pred. No.: 3,06e-224 Length: 547
 Score: 223.00 Matches: 223
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 71.02% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992A-21 (1-994) x NITP_HUMAN (1-547)
 QY 3 CACCCAGTTCCTCCAGATGTTGGTATGCTGCAAAAGACATATGAAATATATGCA 62
 DB |||||||
 DB 156 HSPROVALALAPROGINETPHEGLYTYRALAGLYLSTLHISMETGLUYSTYGLY 175
 QY 63 ACAAAATATGACACCTTTGCAAAAATTTGATGGAATAATCATAACTTCATTAATAC 122
 DB |||||||
 DB 176 ThlylIleGIunIshPheAlaLysIleGIlyTrpLysAshIshSValaAsn 195
 QY 123 CCCTATTCCAGTTCACAGATGTAATACAGTTTGTAGTAAGTATGCAATTAAGAAGTT 182
 DB |||||||
 DB 196 PrrYrSerGIshPheGlnsPgluYrSerLeuAspLlnValMetLaserLysGluVal 215
 QY 183 TTTGATTTTTCAGTATCTTACATGTTTGCACCTTCAGATGCTGCAGCAGCAATT 242
 DB |||||||
 DB 216 PheAsPheuehrIleleuGlnCysProthnSerAspLylAlaAlaIle 235
 QY 243 TTGGCAGTGAACATTTGTACAGAAAGTATGGCTGCATCCAAAGCTGTGGAATTTTG 302
 DB |||||||
 DB 236 LeuAlaSerGIunIshPheValGlnLysTYGLYleuGlnSerLysAlaValGIuIleLeu 255
 QY 303 GCACAGAAATGATGATGCTGTTGCCAGCTGCTTTGAGAAAGCAATTAATAAGT 362
 DB |||||||
 DB 256 AlaGlnIleuLkethrIshPleuProSerSerPheGluIuLysSerIleIleLysMet 275
 QY 363 GTTGCTTTCATATGAGTAAGAAGCTGCAAGAAATGCTATGAGAAATCTGGCTGACA 422
 DB |||||||
 DB 276 ValGIlyPheAsPmetSerLysGluAlaIleArgLysCysTYGLYSerGIlyLeuThr 295
 QY 423 CCAATGATATTTAGCTATATAGAACTTCAGATGCTTTTCTACAAAGACCTTTACT 482
 DB |||||||
 DB 296 PrrAsnAspIleAspValIleGIuLshIshAspCysPheSerThrAsnGIuLeuLeuThr 315
 QY 483 TATGAGCACTGGAGCTGCTGCAGAGGCAAGGCAAGCTGCTGATAGAGAGAT 542
 DB |||||||
 DB 316 TYGIAlaLeuGIlyLeuCysPrrGluGIlyGlnGIlyAlaThrLeuValAspArgLysAsp 335
 QY 543 AATACATATGAGCAAGAAAGTGGCTCATAAATCCATAGTGGTGAATTTCAAGGACAC 602
 DB |||||||
 DB 336 AsnThrTYGLYGIlyLysTrpValIleAsnProSerGIlyGIlyLeuIleSerLysGlnHis 355
 QY 603 CCACATGAGGCGCTACAGAGCTTGCTGCTCAGTGTGCAAGAACTGCTGAGAGGGGAA 662

```

Db 356 ProleuGlyAlaThrGlyLeuAlaGlnCysAlaGlnLeuGlyTrpGlnLeuAlaGlyGlu 375
Oy 663 GCCGGAAGA 671
Db 376 AlaGlyLys 378

RESULT 2
NLTP_MOUSE STANDARD: PRT: 547 AA.
AC p32020:
DT 01-JUN-1993 (Rel. 26, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-TP)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX).
GN SCP2 OR SCP-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93154580; PubMed=8428655;
RA Seedorf U., Raabe M., Assmann G.;
RT *Cloning, expression and sequences of mouse sterol-carrier protein-x-
RT encoding cDNAs and a related pseudogene.*;
RL Gene 123:165-172(1993).
RN [2]
RP SEQUENCE OF 405-547 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=91236770; PubMed=1709640;
RA Monocchi D.T., Pastuszyn A., Scallan T.J.;
RT *cDNA sequence and bacterial expression of mouse liver sterol carrier
RT protein-2.*;
RL J. Biol. Chem. 266:9885-9892(1991).
RN [3]
RP SEQUENCE OF 1-23 FROM N.A.
RC STRAIN=129/SV; TISSUE=Liver;
RX MEDLINE=96354847; PubMed=8751375;
RA Raabe M., Seedorf U., Hamelster H., Ellinghaus P., Assmann G.;
RT *Structure and chromosomal assignment of the murine sterol carrier
RT protein 2 gene (Scp2) and two related pseudogenes by in situ
RT hybridization.*;
RL Cytogenet. Cell Genet. 73:279-281(1996).
CC -1- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
CC PHOSPHOLIPIDS, CHOLESTEROL AND GLANGLIOSIDES BETWEEN MEMBRANES. MAY
CC PLAY A ROLE IN REGULATING STEROIDOGENESIS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDOGENIC TISSUES. SCP-X
CC RESIDES IN THE PEROXISOME.
CC -1- ALTERNATIVE PRODUCTS: 2 Isoforms; SCPx (shown here) and SCP2; are
CC produced by alternative initiation.
CC -1- TISSUE SPECIFICITY: PRESENT AT LOW LEVELS IN ALL TISSUES
CC EXAMINED BUT EXPRESSED PREDOMINANTLY IN THE LIVER.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE THIOLEASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M91458; AAA40098.1; -
CC EMBL: M62361; AAA40099.1; -
CC EMBL: X91150; CAA62592.1; -
CC PIR: A40015; A40015.
CC PIR: J00157; J00157.

```

```

DR MGD; MGI:98254; Scp2.
DR InterPro; IPR003033; SCP2.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00108; thiolase_1.
DR Pfam; PF02036; SCP2; 1.
DR Pfam; PF02803; thiolase_C; 1.
DR PROSITE; PS00098; THIOLEASE_1; 1.
DR PROSITE; PS00737; THIOLEASE_2; 1.
DR PROSITE; PS00099; THIOLEASE_3; FALSE NEG.
DR PROSITE; PS00342; MICROBODIES_CTER; 1.
DR Lipid-binding; Transporc; Mitochondrion; Peroxisome; Transit peptide;
DR Alternative Initiation.
FT CHAIN 1 547
FT TRANSIT 405 424 NONSPECIFIC LIPID-TRANSFER PROTEIN, SCPX
FT CHAIN 425 547 ISOFORM.
FT FT MITOCHONDRION (POTENTIAL).
FT ACT_SITE 405 405 NONSPECIFIC LIPID-TRANSFER PROTEIN, SCP2
FT SITE 94 94 FOR SCP2 ISOFORM.
FT SITE 494 494 SUBSTRATE BINDING (BY SIMILARITY).
FT SITE 494 494 ESSENTIAL FOR TRANSPORT OF LIPIDS (BY
FT SITE 545 547 SIMILARITY).
FT CONFLICT 478 478 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 547 AA; 59158 MW; BAB02F986D16B18 CRC64;
V-> A (IN REF. 2).

Alignment Scores:
Pred. No.: 1,26e-43 Length: 547
Score: 50.00 Matches: 50
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.92% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-21 (1-994) x NLTP_MOUSE (1-547)
Oy 522 ACCTGCTGTATGAGAGATATACATATGAGAGAAAGGCTCATTAATCTAGTGGT 581
Db 329 ThrleuValAspArgGlyAspAsnThrTyrlGlyLysTrpValIleAsnProserGly 348
Oy 582 GGACTGATTCGAAAGGACACCCACTAGCCGCTACAGCTCTTCTCACTGCGACAACTC 641
Db 349 GlyleuIleSerLysGlyLysProleuGlyAlaThrGlyLeuAlaGlnCysAlaGlnLeu 368
Oy 642 TGCTGCGAGCTGAGAGGGAAGCCGGAAGA 671
Db 369 CysTrpGlnLeuArgGlyGluAlaGlyLys 378

RESULT 3
NLTP_MOUSE STANDARD: PRT: 547 AA.
AC p31915: 063383;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nonspecific lipid-transfer protein, mitochondrial precursor (NSL-TP)
DE (Sterol carrier protein 2) (SCP-2) (Sterol carrier protein X) (SCP-X)
DE (SCPX).
GN SCP2 OR SCP-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=91093192; PubMed=1985920;
RA Seedorf U., Assmann G.;
RT *Cloning, expression, and nucleotide sequence of rat liver sterol
RT carrier protein 2 cDNAs.*;
RL J. Biol. Chem. 266:630-636(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91239563; PubMed=2034675;

```

RA Mori T., Tsukamoto T., Mori H., Tashiro Y., Fujiki Y.;
 RT "Molecular cloning and deduced amino acid sequence of nonspecific
 RT lipid transfer protein (sterol carrier protein 2) of rat liver: a
 RT higher molecular mass (60 kDa) protein contains the primary sequence
 RT of nonspecific lipid transfer protein as its C-terminal part.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4338-4342(1991).
 RP [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92007881; PubMed=1915369;
 RA Ossendrop B.C., van Heusden G.P.H., de Beer A.L.J., Bos K.,
 RA Schouten G.L., Wirtz K.W.A.;
 RT "Identification of the cDNA clone which encodes the 58-kDa protein
 RT containing the amino acid sequence of rat liver non-specific lipid-
 RT transfer protein (sterol carrier protein 2). Homology with rat
 RT peroxisomal and mitochondrial 3-oxoacyl-CoA thiolases.";
 RL Eur. J. Biochem. 201:233-239(1991).
 RP [4]
 RP SEQUENCE OF 33-547 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=90241231; PubMed=2334427;
 RA Ossendrop B.C., van Heusden G.P.H., Wirtz K.W.A.;
 RT "The amino acid sequence of rat liver non-specific lipid transfer
 RT protein (sterol carrier protein 2) is present in a high molecular
 RT weight protein: evidence from cDNA analysis.";
 RL Biochem. Biophys. Res. Commun. 168:631-636(1990).
 RN [5]
 RN SEQUENCE OF 275-547 FROM N.A.
 RC STRAIN=CD Charles River; TISSUE=Liver;
 RX MEDLINE=90253610; PubMed=2340090;
 RA Billheimer J.T., Strehl L.L., Strauss J.F. III,
 RA Davis L.G.;
 RT "Characterization of a cDNA encoding rat sterol carrier protein2";
 RL DNA Cell Biol. 9:159-165(1990).
 RN [6]
 RN SEQUENCE OF 425-547.
 RC TISSUE=Liver;
 RX MEDLINE=88007528; PubMed=3115977;
 RA Pastuszyn A., Noland B.J., Bazan J.F., Fletterick R.J., Scallen T.J.;
 RT "Primary sequence and structural analysis of sterol carrier protein 2
 RT from rat liver: homology with immunoglobulins.";
 RL J. Biol. Chem. 262:13219-13227(1987).
 RN [7]
 RN SEQUENCE OF 425-547.
 RC TISSUE=Liver;
 RX MEDLINE=88280812; PubMed=3395344;
 RA Morris H.R., Larsen B.S., Billheimer J.T.;
 RT "A mass spectrometric study of the structure of sterol carrier
 RT protein SCP2 from rat liver.";
 RL Biochem. Biophys. Res. Commun. 154:476-482(1988).
 CC -I- FUNCTION: MEDIATES IN VITRO THE TRANSFER OF ALL COMMON
 CC PHOSPHOLIPIDS, CHOLESTEROL AND GANGLIOSIDES BETWEEN MEMBRANES. MAY
 CC PLAY A ROLE IN REGULATING STEROIDGENESIS.
 CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC IN THE LIVER AND ALSO ASSOCIATED
 CC WITH MITOCHONDRIA ESPECIALLY IN STEROIDGENIC TISSUES. SCP-X
 CC RESIDES IN THE PEROXISOME.
 CC -I- ALTERNATIVE PRODUCTS: 2 isoforms: SCPx (shown here) and SCP2; are
 CC produced by alternative initiation.
 CC -I- TISSUE SPECIFICITY: LIVER > INTESTINE > BRAIN > LUNG, COLON,
 CC STOMACH, SPLEEN, KIDNEY, HEART, AND OVARY.
 CC -I- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE THIOLASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M57453; AAA42122.1; -;
 DR EMBL; M58287; AAA41726.1; -;
 DR EMBL; M34728; AAA42120.1; -;
 DR EMBL; X60654; CAA43060.1; ALT_INIT.
 DR EMBL; X60654; CAA43061.1; -;
 DR PIR; A27661; A27661.
 DR PIR; A29366; A29366.
 DR PIR; A34584; A34584.
 DR PIR; A34635; A34635.
 DR PIR; A39054; A39054.
 DR InterPro: IPR003033; SCP2.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF00108; thiolase, 1.
 DR Pfam: PF02036; SCP2; 1.
 DR Pfam: PF02803; thiolase_C; 1.
 DR PROSITE: PS00099; THIOLASE_3; FALSE_NEG.
 DR PROSITE: PS00098; THIOLASE_1; 1.
 DR PROSITE: PS00342; MICRORODIES_CTER; 1.
 DR PROSITE: PS00737; THIOLASE_2; 1.
 DR Lipid-binding; Transport; Mitochondrion; Peroxisome; Transit peptide;
 KW Alternative initiation.
 FT CHAIN 1 547
 FT TRANSIT 405 424
 FT CHAIN 425 547
 FT INIT_MET 405 405
 FT ACT_SITE 94 94
 FT SITE 494 494
 FT SITE 545 547
 FT CONFLICT 12 12
 FT CONFLICT 50 50
 FT CONFLICT 265 265
 FT CONFLICT 427 428
 FT CONFLICT 436 436
 FT CONFLICT 446 446
 FT CONFLICT 450 450
 FT CONFLICT 488 488
 FT CONFLICT 516 516
 FT CONFLICT 526 526
 FT CONFLICT 537 537
 FT CONFLICT 543 543
 FT CONFLICT 547 AA; 58813 MW; DDD1B435D2DC6AFB CRC64;
 SQ SEQUENCE

Alignment Scores:
 Pred. No.: 1.39e-42 Length: 547
 Score: 49.00 Matches: 49
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.61% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992A-21 (1-994) x NLRP_RAT (1-547)
 QY 525 CIGCTTGATAGAGGAGATATATGAGAGAAAGTGGCTATTAATCTTGCTGCA 584
 DB 330 LeuValAspArgGlyAspAsnThrTyrGlyGlyTyrValIleAsnProserGlyGly 349
 QY 585 CTGATTCAAGAGGAGACACCTAGCGCTGACAGGCTGCAGTGCGACAGCTGCG 644
 DB 350 LeuIleSerTyrGlyHisProLeuGlyIleAlaGlnGlyLeuAlaGlnGlyLeuGly 369
 QY 645 TGCAGCTGAGAGGAGAACCGGAAAA 671
 DB 370 TrpGlnLeuArgGlyGlnIleGlyLys 378
 RESULT 4
 NLRP_CHICK STANDARD: PRT: 547 AA.
 ID NLRP_CHICK
 AC 007598;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

OS Apterix australis (Brown kiwi).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Palaeognathae; Apterygiformes; Apteryidae;
 OC Apteryx.
 RN NCBI_TaxID=8822;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA MEDLINE=95331397; PubMed=7541757;
 RA Hjelmgvist L., Metts M., Persson H., Hoeog J.-O., McLennan J.,
 RA Joernvall H.;
 RT "Alcohol dehydrogenase of class I: kiwi liver enzyme, parallel
 RT evolution in separate vertebrate lines, and correlation with 12S rRNA
 RT patterns.";
 RL FEBS Lett. 367:306-310(1995).
 CC -1- CATALYTIC ACTIVITY: An alcohol + NAD(+) = an aldehyde or ketone +
 CC NADH.
 CC -1- COFACTOR: ZINC.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S78778; AAC60755.2; -.
 DR HSSP: P00327; 20HX.
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; adh_zinc; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; NAD; Acetylation.
 FT INIT_MET 0
 FT MOD_RES 1
 FT METAL 46
 FT METAL 46
 FT METAL 67
 FT METAL 97
 FT METAL 100
 FT METAL 103
 FT METAL 111
 FT METAL 111
 FT METAL 174
 SQ SEQUENCE 374 AA; 39500 MW; 4D1BD387891F63E6 CRC64;
 Alignment Scores:
 Pred. No.: 9.49 Length: 374
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.58% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992a-21 (1-994) x ADH1-APPAU (1-374)
 QY 247 GCCAAATTTGCTGTCGACGACCA 224
 DB 158 AlalysllleAlaAlaAlaAlaPro 165
 RESULT 7
 VANS_ENTFC STANDARD: PRT; 384 AA.
 AC 006240:
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sensor protein vans (EC 2.7.3.-) (Vancomycin resistance protein vans)
 DE (Vancomycin histidine protein kinase).
 GN VANS.
 OS Enterococcus faecium (Streptococcus faecium).
 OS Plasmid pIP816.

OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BM4147; TRANSPOSON=TN1546;
 RX MEDLINE=93106944; PubMed=8380148;
 RA Arthur M., Molinas C., Delpardieu F., Courvalin P.;
 RT "Characterization of Tn1546, a Tn3-related transposon conferring
 RT glycopeptide resistance by synthesis of depsipeptide peptidoglycan
 RT precursors in Enterococcus faecium BM4147.";
 RL J. Bacteriol. 175:117-127(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BM4147;
 RX MEDLINE=92210502; PubMed=1556077;
 RA Arthur M., Molinas C., Courvalin P.;
 RT "The Vans-Vanr two-component regulatory system controls synthesis of
 RT depsipeptide peptidoglycan precursors in Enterococcus faecium
 RT BM4147.";
 RL J. Bacteriol. 174:2582-2591(1992).
 CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM VANS/VANR.
 CC ACTIVATES THE TRANSCRIPTION OF VANH, VANA AND VANX IN RESPONSE TO
 CC VANCOMYCIN WHICH RESULTS IN VANCOMYCIN RESISTANCE. VANS ACTIVATES
 CC VANR BY PHOSPHORYLATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M9297; AAA65964.1; -.
 DR EMBL: M68910; AAA24788.1; -.
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR InterPro: IPR003661; His_KINDA.
 DR Pfam: PF00512; signal; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR SMART: SM00387; HATPase_C; 1.
 DR SMART: SM00388; HATPase_C; 1.
 DR PROSITE: PS50109; HIS_KIN; 1.
 KW Sensory transduction; His_KIN; 1.
 KW Antibiotic resistance; Phosphorylation; Plasmid.
 FT TRANSMEM 21
 FT TRANSMEM 41
 FT TRANSMEM 76
 FT DOMAIN 96
 FT MOD_RES 161
 FT MOD_RES 164
 SQ SEQUENCE 384 AA; 43915 MW; D69099844EF37C35A CRC64;
 Alignment Scores:
 Pred. No.: 9.45 Length: 384
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.55% Indels: 0
 DB: 1 Gaps: 0
 US-09-835-992a-21 (1-994) x VANS_ENTFC (1-384)
 QY 58 ATGGAACAAATTTGACACTTG 81
 DB 130 MetGluglnLysLeuasnThrnu 137
 RESULT 8
 Y666_METJA STANDARD: PRT; 398 AA.
 AC 058080:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)


```

OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=cv_SRI;
RX MEDLINE=98088034; PubMed=9426628;
RA Kruse E., Mock H.-P., Grimm B.;
RT "Isolation and characterisation of tobacco (Nicotiana tabacum) cDNA
RT clones encoding proteins involved in magnesium chelation into
RT protoporphyrin IX."
RL Plant Mol. Biol. 35:1053-1056(1997).
CC -I FUNCTION: Involved in chlorophyll IX to yield Mg-protoporphyrin
CC magnesium ion into protoporphyrin IX
CC IX.
CC -I PATHWAY: Chlorophyll biosynthesis.
CC -I SUBCELLULAR LOCATION: Chloroplast stroma.
CC -I TISSUE SPECIFICITY: Strongly expressed in young leaves and to a
CC lesser extent in mature leaves.
CC -I SIMILARITY: BELONGS TO THE MG-CHELATASE SUBUNITS D/I FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
-----
CC
CC DR EMBL; AF014053; AAB97153.1; -.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR00523; Mg_chelatase_chi1.
CC DR Pfam; PF01078; Mg_chelatase_1.
CC DR SMART; SMO0382; AAA; 1.
CC KM Photosynthesis; Chlorophyll biosynthesis; Chloroplast;
CC KW Transit peptide; ATP-binding.
CC FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC FT CHAIN ? 426 MAGNESIUM-CHELATASE SUBUNIT CHL1.
CC FT NP_BIND 121 ATP (POTENTIAL).
CC SQ SEQUENCE 426 AA; 46627 MW; 42B7A6F04E1C3274 CRC64;
-----
Alignment Scores:
Pred. NO.:          9.28      Length:    426
Score:              8.00      Matches:     8
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:        2.55%     Indels:      0
DB:                  1       Gaps:         0
US-09-835-992A-21 (1-994) x CHL1_TOBAC (1-426)
QY 228 CCTGCAGCACGCAATTTGGCCACT 251
      |||||||
Db 11 AlaAlaAlaAlaIleLeuAlaser 18
-----
RESULT 11
POLG_WMV2F STANDARD; PRT; 455 AA.
AC Q89251;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Genome polypeptide [Contains: Nuclear inclusion protein B (NIB-B) (NIB)
DE (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP)]
DE (Fragment).
OS Maternelon mosaic virus II (Isolate Tonga) (Vanilla necrosis
OS potyvirus) (VNV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
CN 11
RN NCBI_TaxID=148359;
RP SEQUENCE FROM N.A..
RX MEDLINE=93228466; PubMed=8470960;
Wang Y.Y., Beck D.L., Gardner R.C., Pearson M.N.;
"Nucleotide sequence, serology and symptomatology suggest that vanilla

```

```

RT      necrosis polyvirus is a strain of watermelon mosaic virus II." ;
RL      Arch. Virol. 129:93-103(1993)
CC      -I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC      (RNA)(N).
CC      -I- PWM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC      POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC      PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC      INDIVIDUAL PROTEINS.
CC      -I- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC      -----
CC      CC      This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC      CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC      the European Bioinformatics Institute. There are no restrictions on its
CC      CC      use by non-profit institutions as long as its content is in no way
CC      CC      modified and this statement is not removed. Usage by and for commercial
CC      CC      entities requires a license agreement (see http://www.isb.ch/announce/
CC      CC      or send an email to license@isb-sib.ch).
CC      -----
CC      DR      EMBL: L22907: AAAA8497.2; ALT_INIT.
CC      DR      InterPro: IPR001592; Poty_coat.
CC      DR      InterPro: IPR001205; RNA_pol_P3D.
CC      DR      Pfam: PF00680; RNA_dep_RNA_pol_1.
CC      DR      Pfam: PF00767; Poty_coat; 1.
CC      KW      Transferrase; RNA-directed RNA polymerase; Coat protein; Polyprotein.
CC      FT      NON_TER      1      1
CC      FT      CHAIN      <1      172      NUCLEAR INCLUSION PROTEIN B.
CC      FT      CHAIN      173      455      COAT PROTEIN.
CC      SQ      SEQUENCE      455 AA; 51607 MW; 937AFDEAE9A52EC1 CRC64;
CC      -----
CC      Alignment Scores:
CC      Pred. No.:      9.18      Length:      455
CC      Score:      8.00      Matches:      8
CC      Best Similarity:      100.00%      Conservative:      0
CC      Best Local Similarity:      100.00%      Mismatches:      0
CC      Query Match:      2.58%      Indels:      0
CC      DB:      1      Gaps:      0
CC      -----
US-09-835-992A-21 (1-994) x POLG_RMV2T (1-455)
QY      782 AAGGAACGTGGCGCTTCGCGGAA 759
        |||||||
DB      122 LysgltleuAlaAlaIasercglyls 129
        |||||||
RESULT 12
ID      Y456_CHLTR      STANDARD;      PRT;      1005 AA.
AC      084462;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Protein CT456 precursor.
GN      CT456.
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RA      NCBI_Taxid=813;
RM      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D/UW-3/Cx;
RX      MEDLINE=99000809; PubMed=9784136;
RA      Stephens R.S., Kalman S., Lammel C.J., Fan J., Koonin R., Aravind L.,
RA      Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA      Davis R.W.;
RT      "Genome sequence of an obligate intracellular pathogen of humans:
RT      Chlamydia trachomatis." ;
RL      Science 282:754-759(1998).
CC      -I- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC041
CC      FAMILY.
CC      -----
CC      CC      This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC      CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC      the European Bioinformatics Institute. There are no restrictions on its
CC      CC      use by non-profit institutions as long as its content is in no way
CC      CC      modified and this statement is not removed. Usage by and for commercial

```


CC entitles requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).

CC -----

DR EMBL: AB001319: AAC68056.1: -

KW PHCI-2DPAGE: 084462: -

FT SIGNAL: Complete proteome.

FT CHAIN 1 40 POTENTIAL.

FT SIGNAL 41 1005 PROTEIN CT456.

SO SEQUENCE 1005 AA; 102131 MW; EC47EC389851CDIE CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Gaps:
Score: 8.01	1005	8	0	100.00%	0	0	0
Percent Similarity: 100.00%							
Best Local Similarity: 100.00%							
Query Match: 2.55%							
DB: 1							

US-09-835-992a-21 (1-994) x Y456.CHLTR (1-1005)

Oy 216 ACTTCAGATGTCGTCACACAGCA 239

DB 285 ThSeraSpCjYAlaAlaAlaAla 292

RESULT 13

POLG_MMV2U STANDARD; PRT; 1016 AA.

ID POLG_MMV2U

AC P18478:

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Genome polyprotein (Contains: Nuclear inclusion protein A (NI-A) (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP) (Fragment).

DE Watermelon mosaic virus II (Isolate USA).

OS Watermelon mosaic virus II (Isolate USA).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae; Polyvirus.

OX NCBI_TaxID=148360;

RN [1]

RP SEQUENCE FROM N.A.

RA Slightom J.L.;

RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE OF 736-1016 FROM N.A.

RX MEDLINE-90324931; PubMed-2374006;

RT "Quemada H., Steu L.C., Stemieniak D.R., Consalves D., Slightom J.L.; Watermelon mosaic virus II and zucchini yellow mosaic virus: cloning of 3'-terminal regions, nucleotide sequences, and phylogenetic comparisons."

RT J. Gen. Virol. 71:1451-1460(1990).

RL J. Gen. Virol. 71:1451-1460(1990).

CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY. further restricted by preferences for the amino acids in P6 - P1' that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-Gln-(Ser or Gly) for the enzyme from tobacco etch virus. The natural substrate is the viral polyprotein, but other proteins and oligopeptides containing the appropriate consensus sequence are also cleaved.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA) (N).

CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.

CC -1- BIOCHEMISTRY: The gene for the coat protein is introduced by genetic manipulation and expressed in squash so as to obtain virus resistant plants.

CC -1- MISCELLANEOUS: READTHROUGH OF A TERMINATOR CODON TGA OCCURS BETWEEN RESIDUES 267-ALA AND 268-GLN.

CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.

CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial CC entitles requires a license agreement (See <http://www.isb-slb.ch/announce/> or send an email to license@isb-slb.ch).

CC -----

DR EMBL: D13913; BAA03009.1: -

DR PIR: J00498; J00498.

DR MEROPS: C04.003: -

DR InterPro: IPR001730; Peptidase_C4.

DR InterPro: IPR001592; Poly_coat.

DR InterPro: IPR001205; RNA_pol_P3D.

DR Pfam: PF00680; RNA_dep_RNA_pol; 1.

DR Pfam: PF00767; Poly_coat; 1.

DR Pfam: PF00863; Peptidase_C4; 1.

KW Hydroxylase: Transferase; Thiol protease; RNA-directed RNA polymerase; Coat protein; Polyprotein; Genetically modified food.

FT NON_TER 1 1

FT CHAIN <1 219 NUCLEAR INCLUSION PROTEIN A.

FT CHAIN 220 736 NUCLEAR INCLUSION PROTEIN B.

FT CHAIN 737 1016 COAT PROTEIN.

FT SEQUENCE 1016 AA; 115377 MW; 8179821C4AFAD5D CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Best Local Similarity:	Mismatches:	Indels:	Gaps:
Score: 7.99	1016	8	0	100.00%	0	0	0
Percent Similarity: 100.00%							
Best Local Similarity: 100.00%							
Query Match: 2.58%							
DB: 1							

US-09-835-992a-21 (1-994) x POLG_MMV2U (1-1016)

Oy 782 AAGCACTGGCGCTTCGGGANA 759

DB 683 LysGluLeuAlaAlaLaserGlyLys 690

RESULT 14

PAND_ECOLI STANDARD; PRT; 126 AA.

ID PAND_ECOLI

AC P31664:

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Aspartate 1-decarboxylase precursor (EC 4.1.1.11) (Aspartate alpha-decarboxylase).

DE PAND OR B0131 OR 20142 OR ECS0135.

OS Escherichia coli, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562, 83334;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-K12 / W3110;

RL Merkel W.K., Nichols B.P.;

RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / M61655;

RX MEDLINE-97426617; PubMed-9278503;

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE-21074935; PubMed-11206551;

RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsuda E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [5]
 RP SEQUENCE OF 1-5 AND 25-29, AND CHARACTERIZATION.
 RX MEDLINE=97270610; PubMed=9169598;
 RA Ramjee M.K., Genschel U., Abell C., Smith A.G.;
 RT "Escherichia coli L-aspartate-alpha-decarboxylase: preprotein
 RT processing and observation of reaction intermediates by electrospray
 RT mass spectrometry.";
 RL Biochem. J. 323:661-669(1997).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=98206295; PubMed=9546220;
 RA Albert A., Dhanraj V., Genschel U., Khan G., Ramjee M.K., Pulido R.,
 RA Sibanda B.L., von Delft F., Wilty M., Blundell T.L., Smith A.G.,
 RA Abell C.;
 RT "Crystal structure of aspartate decarboxylase at 2.2-A resolution
 RT provides evidence for an ester in protein self-processing.";
 RL Nat. Struct. Biol. 5:289-293(1998).
 CC -1- CATALYTIC ACTIVITY: L-aspartate = beta-alanine + CO(2).
 CC -1- COFACTOR: Pyruvoyl group.
 CC -1- PATHWAY: Pantothenate biosynthesis; second branch.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE PAND FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L17086; AAA24274.1; -;
 DR EMBL; AE000122; AAC73242.1; -;
 DR EMBL; AE005189; AAG54435.1; -;
 DR EMBL; AP002550; BAB33558.1; -;
 DR PDB; 1AW8; 29-APR-98.
 DR Ecogene; Egl1747; pand.
 DR InterPro; IPR003190; Asp.decarbox.
 DR Pfam; PF02261; Asp.decarbox; 1.
 DR TIGRFAMs; TIGR00223; pand; 1.
 DR Pantothenate biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
 KM 3D-structure; Complete proteome.
 FT CHAIN 1 24 ASPARTATE 1-DECARBOXYLASE BETA CHAIN.
 FT MOD_RES 25 25 CONVERTED TO A PYRUVOYL GROUP.
 FT SDOURCE 126 AA; 13834 MW; E3169F5C2BD5D25 CRC64;
 SQ

Alignment Scores:

Pred. No.: 127
 Score: 7.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 2.23%
 DB: 1 Gaps: 0

US-09-835-992a-21 (1-994) x PAND_ECOLI (1-126)
 Oy 647 GCAGCTGAGAGCGGAAGCCGG 667
 Db 61 AAlaAGlunAGGlySerArg 67
 RESULT 15
 ID PAND_SALTY STANDARD; PRT; 126 AA.
 AC O8XG14;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aspartate 1-decarboxylase precursor (EC 4.1.1.11) (Aspartate alpha-
 DE decarboxylase).
 GN PAND OR STM0180 OR STY0198.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. typhimurium; STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Speleth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Conerton P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagers K.,
 RA Krogsh A., Larsen T.S., Leather S., Moule S., O'Goira P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT *enterica* serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 CC -1- CATALYTIC ACTIVITY: L-aspartate = beta-alanine + CO(2).
 CC -1- COFACTOR: Pyruvoyl group (By similarity).
 CC -1- PATHWAY: Pantothenate biosynthesis; second branch.
 CC -1- SIMILARITY: BELONGS TO THE PAND FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE008702; AAL19144.1; -;
 DR EMBL; AL627265; CAD01334.1; -;
 DR StGene; SG72727; pand.
 DR InterPro; IPR003190; Asp.decarbox.
 DR Pfam; PF02261; Asp.decarbox; 1.
 DR TIGRFAMs; TIGR00223; pand; 1.
 DR Pantothenate biosynthesis; Lyase; Decarboxylase; Pyruvate; Zymogen;
 KM Complete proteome.
 FT CHAIN 1 24 ASPARTATE 1-DECARBOXYLASE BETA CHAIN
 FT MOD_RES 25 25 CONVERTED TO A PYRUVOYL GROUP.
 FT SDOURCE 126 AA; 13834 MW; E3169F5C2BD5D25 CRC64;
 SQ

FT MOD_RES 25 25 (BY SIMILARITY)
FT CONVERTED TO A PYRUVYL GROUP
SQ SEQUENCE 126 AA; 13887 MW; 3A9358C0012AA80 CRC64;
Alignment Scores:
Pred. No.: 127
Score: 7.00 Length: 126
Percent Similarity: 100.00% Matches: 7
Best Local Similarity: 100.00% Conservative: 0
Query Match: 2.23% Mismatches: 0
DB: 1 Indels: 0
Gaps: 0
US-09-835-992a-21 (1-994) x PAND_SALTY (1-126)
QY 647 GCAGCTGAGAGCGGAACCCG 667
|||||
Db 61 Alalaeqluargllyserarg 67
Search completed: January 14, 2003, 17:29:17
Job time: 21.349 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: January 14, 2003, 17:20:59 ; Search time 69.9914 Seconds
(without alignments)
5852.461 Million cell updates/sec

Title: US-09-835-992A-21

Perfect score: 314
Sequence: 1 ctcaccacgtgtctctccag.....ccaccctggggtgtgtggat 994

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Xgapop 60.0 , Xgapext 60.0
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL-frame+ n2p.model -DEV-xlp
-O/cgn2.1/USPTO/US09835992/runat.14012003.161526.4027/app.query.fasta_1.3932
-DB-SPTREMBL-21 -OBMT-fastan -SUFFIX-olig.rspt -MINMATCH-0 -1-LOOEXT-0
-LOOEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-oligo -TRANS-human40.cdi
-LIST-45 -DOCALLIGN-200 -THR-SCORE-quality -THR_MIN-1 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-US09835992-ecgn.1.1.346-ecgnat.14012003.161526.4027 -NCPU-6 -ICPU-3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -LONGLOG -DEV-TIMEOUT-120
-NARN-TIMEOUT-30 -THREADS-1 -XGAPOP-60 -XGAPEXT-60 -FGAPOP-6 -FGAPEXT-7
-YGAPOP-60 -YGAEXT-60 -DELOP-6 -DELEXT-7

Database :

- 1: SPTREMBL_21:*
- 2: SP_Archea:*
- 3: SP_Bacteria:*
- 4: SP_Fungi:*
- 5: SP_Human:*
- 6: SP_Invertebrate:*
- 7: SP_Mhc:*
- 8: SP_Organellae:*
- 9: SP_Phage:*
- 10: SP_Plant:*
- 11: SP_Rodent:*
- 12: SP_Virus:*
- 13: SP_Vertebrate:*
- 14: SP_Unclassified:*
- 15: SP_RVirus:*
- 16: SP_Bacteriophage:*
- 17: SP_Archaeo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	50	15.9	547 11 Q9DBM7	Q9dbm7 mus musculu

2	49	15.6	405 11 Q9QW19	Q9qwi9 rattus sp.
3	49	15.6	547 6 Q62742	Q62742 oryctolagus
4	30	9.6	378 5 Q9VJ44	Q9vj44 drosophila
5	407	9.6	407 5 Q95TG8	Q95tg8 drosophila
6	30	9.6	414 5 Q24507	Q24507 drosophila
7	30	9.6	543 5 Q24506	Q24506 drosophila
8	30	9.6	544 5 Q9VJ43	Q9vj43 drosophila
9	24	7.6	401 16 Q33332	Q33332 mycobacteri
10	22	7.0	412 5 Q02353	Q02353 caenorhabdit
11	21	6.7	400 16 Q8VJDS	Q8vjds mycobacteri
12	20	6.4	132 2 Q9E205	Q9e205 streptomyce
13	19	6.1	152 6 Q28224	Q28224 sus scrofa
14	19	3.2	387 17 Q97VJ9	Q97vj9 sulfolobus
15	16	3.2	387 17 Q96X09	Q96x09 sulfolobus
16	10	3.2	246 2 Q47870	Q47870 erysipeloth
17	9	2.9	351 17 Q9ZU28	Q9zu28 arabidopsis
18	9	2.9	383 17 Q26884	Q26884 methanobact
19	9	2.9	390 17 Q97CN3	Q97cn3 thermoplasm
20	9	2.9	391 17 Q9HM39	Q9hm39 thermoplasm
21	9	2.5	510 2 Q9RA39	Q9ra39 trichodesmi
22	8	2.5	151 8 Q9MTP2	Q9mtp2 oenothera h
23	8	2.5	224 11 Q9D314	Q9d314 mus musculu
24	8	2.5	231 5 Q17964	Q17964 caenorhabdit
25	8	2.5	267 16 Q50527	Q50527 streptomyce
26	8	2.5	299 17 Q82TV2	Q82tv2 pyrobaculum
27	8	2.5	308 9 Q38577	Q38577 bacteriophag
28	8	2.6	314 10 Q9SZB9	Q9szb9 arabidopsis
29	8	2.5	328 5 Q9N965	Q9n965 leishmania
30	8	2.5	374 2 Q938V4	Q938v4 bradyrhizob
31	8	2.5	379 16 Q9HZ07	Q9hzo7 pseudomonas
32	8	2.5	383 17 Q9H028	Q9h028 halobacteri
33	8	2.5	387 16 P96838	P96838 mycobacteri
34	8	2.5	387 17 Q29811	Q29811 archaeoglob
35	8	2.5	388 17 Q29294	Q29294 archaeoglob
36	8	2.5	391 17 Q30217	Q30217 archaeoglob
37	8	2.5	394 16 Q99W22	Q99w22 staphylococ
38	8	2.5	394 17 Q8TVL1	Q8tv11 methanopyru
39	8	2.5	395 17 Q97VV9	Q97vv9 sulfolobus
40	8	2.5	395 17 Q8ZV25	Q8zv25 pyrobaculum
41	8	2.5	396 17 Q978T8	Q978t8 thermoplasm
42	8	2.5	396 17 Q9H1D2	Q9h1d2 thermoplasm
43	8	2.5	397 2 Q93SM1	Q93sm1 staphylococ
44	8	2.5	399 16 Q8Y261	Q8y261 ralsionia s
45	8	2.5	400 16 Q9ABC2	Q9abc2 caulobacter

ALIGNMENTS

RESULT 1	ID	Q9DBM7	PRELIMINARY;	PRT; 547 AA.
AC	Q9DBM7	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	steroid carrier protein 2, liver.			
GN	SCP2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=LIVER;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,			
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gasteirland T., Glassl C., King B., Kochius H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,			

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gueniche S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuk S.,
 RA Hayashizaki Y.,
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 RN
 RP SEQUENCE FROM N.A.
 RA Strausberg R.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK004860; BAB3623.1; -
 DR EMBL: BC018384; AAH18384.1; -
 DR MGI: 98254; SGP2.
 DR InterPro: IPR003033; SCP2.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF02036; SCP2; 1.
 DR Pfam: PF00108; thiolase; 1.
 DR Pfam: PF02803; thiolase; 1.
 DR PROSITE: PS00098; THIOLEASE_1; 1.
 DR PROSITE: PS00737; THIOLEASE_2; 1.
 SQ SEQUENCE 547 AA; 59125 MW; 37BA2E730D9CB105 CRC64;

Alignment Scores:

Pred. No.:	6.52e-45	Length:	547
Score:	50.00	Matches:	50
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	15.92%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992a-21 (1-994) x Q9QW19 (1-547)

QY 522 ACGGTGTTGATAGAGAGATATACATATGAGAGAAAGTGGTCATTAATCTAGTGT 581
 DB 329 ThrLeuValAspArgGlyAspAsnThrTyrGlyGlySTPValIleAsnProSerGly 348
 QY 582 GCAGTATTCAAAAGGACACCCACTAGCGCTACAGTCTTGCTCAGTGTGCAACTGC 644
 DB 349 GlyLeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGlnLeu 368
 QY 642 TGGCAGCTGAGAGGAGGAGCGGAAAA 671
 DB 369 CysTrpGlnLeuArgGlyGlnAlaGlyLys 378

RESULT 2

Q9QW19 PRELIMINARY; PRT; 405 AA.
 AC Q9QW19;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE 3-oxoacyl-CoA thiolase homolog (Fragment).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93283441; PubMed=8507681;
 RA Ossendorp B.C., Wirtz K.W.,
 RT "The non-specific lipid-transfer protein (sterol carrier protein 2)
 RT and its relationship to peroxisomes.";
 RL Biochimie 75:191-200(1993).
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF00108; thiolase; 1.
 DR Pfam: PF02803; thiolase; 1.
 DR PROSITE: PS00098; THIOLEASE_1; 1.
 DR PROSITE: PS00098; THIOLEASE_2; 1.

DR PROSITE: PS00737; THIOLEASE_2; 1.
 SQ SEQUENCE 405 AA; 43650 MW; 8C6782095D7F628F CRC64;
 Alignment Scores:

Pred. No.:	8.26e-44	Length:	405
Score:	49.00	Matches:	49
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	15.61%	Indels:	0
DB:	11	Gaps:	0

US-09-835-992a-21 (1-994) x Q9QW19 (1-405)

QY 525 CTGGTTGATAGAGAGATATACATATGAGAGAAAGTGGTCATTAATCTAGTGTGA 584
 DB 329 LeuValAspArgGlyAspAsnThrTyrGlyGlySTPValIleAsnProSerGly 348
 QY 585 CTGATTTCAAAAGGACACCCACTAGCGCTACAGTCTTGCTCAGTGTGCAACTGC 644
 DB 349 LeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGlnLeu 368
 QY 645 TGGCAGCTGAGAGGAGGAGCGGAAAA 671
 DB 369 TrpGlnLeuArgGlyGlnAlaGlyLys 377

RESULT 3

ID	062742	PRELIMINARY;	PRT; 547 AA.
AC	062742;		
DT	01-AUG-1998 (TREMBlrel. 07, Created)		
DT	01-AUG-1998 (TREMBlrel. 07, Last sequence update)		
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)		
DE	Sterol carrier protein X.		
GN	SCP2.		
OS	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Weber F.E., Weider M., Boffelli D., Hauser H.;		
RT	"In pre-sterol carrier protein 2 (SCP2) in solution the leader peptide		
RT	1-20 is flexibly disordered and the residues 21-143 adopt the same		
RT	globular fold as in mature SCP2.";		
RL	Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF051897; AAC15422.1; -		
DR	HSSP: P27796; 1PXT.		
DR	InterPro: IPR003033; SCP2.		
DR	InterPro: IPR002155; Thiolase.		
DR	Pfam: PF02036; SCP2; 1.		
DR	Pfam: PF00108; thiolase; 1.		
DR	Pfam: PF02803; thiolase; 1.		
DR	PROSITE: PS00098; THIOLEASE_1; 1.		
DR	PROSITE: PS00737; THIOLEASE_2; 1.		
SQ	SEQUENCE 547 AA; 58903 MW; 32BE7F7DCBF9BCD7 CRC64;		

Alignment Scores:

Pred. No.:	7.97e-44	Length:	547
Score:	49.00	Matches:	49
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	15.61%	Indels:	0
DB:	6	Gaps:	0

US-09-835-992a-21 (1-994) x 062742 (1-547)

QY 525 CTGGTTGATAGAGAGATATACATATGAGAGAAAGTGGTCATTAATCTAGTGTGA 584
 DB 330 LeuValAspArgGlyAspAsnThrTyrGlyGlySTPValIleAsnProSerGly 349
 QY 585 CTGATTTCAAAAGGACACCCACTAGCGCTACAGTCTTGCTCAGTGTGCAACTGC 644
 DB 350 LeuIleSerLysGlyHisProLeuGlyAlaThrGlyLeuAlaGlnCysAlaGlnLeu 369

QY 645 TGGCAGCTGACAGCGCAAGCCGAAAA 671
Db 370 TrpGlnLeuArgGlyGlnAlaGlyLys 378

RESULT 4

09VJ44 PRELIMINARY; PRT; 378 AA.
ID 09VJ44
AC 09VJ44:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE CG17597 protein.
GN CG17597.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RX MEDLINE=20196006; PubMed=10731132;
RC STRAIN=BERKELEY;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs R., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor B.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iegem C.,
RA Jalili M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splet E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003660; AAF53712.1;
DR FLYbase: FBgn0032715; CG17597.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; Thiolase; 1.
DR Pfam: PF02803; Thiolase; 1.
DR PROSITE: PS00098; THIOBASE_1; 1.
DR PROSITE: PS00737; THIOBASE_2; 1.
SQ SEQUENCE 378 AA; 40655 MW; 704CB62E7E6F075E CRC64;

Alignment Scores:

pred. No.: 3,76e-23 Length: 378
Score: 30.00 Matches: 30

Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 9.558
DB: 5
Gaps: 0

US-09-835-992a-21 (1-994) x 09VJ44 (1-378)

QY 570 AATCTAGTGTGAGCTGATTTCAAGGACACCCAGGCGCTACAGTCTTCTCAG 629
Db 308 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaTrnGlyLeuAlaGln 327
QY 630 TGGCAGACTCTGCTGCGACAGCTGAGAGC 659
Db 328 CysAlaGlyLeuGlyCysTrpGlnLeuArgGly 337

RESULT 5

095TG8 PRELIMINARY; PRT; 407 AA.
ID 095TG8
AC 095TG8:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SP06839P.
GN CG17597.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RX MEDLINE=20196006; PubMed=10731132;
RC STRAIN=BERKELEY;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs R., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman J.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor B.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iegem C.,
RA Jalili M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Splet E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL: AE003660; AAF53712.1;
DR FLYbase: FBgn0032715; CG17597.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; Thiolase; 1.
DR Pfam: PF02803; Thiolase; 1.
DR PROSITE: PS00098; THIOBASE_1; UNKNOWN.1.
DR PROSITE: PS00737; THIOBASE_2; UNKNOWN.1.
SQ SEQUENCE 407 AA; 43904 MW; 7BE64B5D68C49D1B CRC64;

Alignment Scores:

pred. No.: 3,73e-23 Length: 407
Score: 30.00 Matches: 30
Percent Similarity: 100.008
Best Local Similarity: 100.008
Query Match: 9.558
DB: 5
Gaps: 0

US-09-835-992a-21 (1-994) x 095TG8 (1-407)

QY 570 AATCTAGTGTGAGCTGATTTCAAGGACACCCAGGCGCTACAGTCTTCTCAG 629
Db 337 AsnProserGlyGlyLeuIleSerLysGlyHisProLeuGlyAlaTrnGlyLeuAlaGln 356
QY 630 TGGCAGACTCTGCTGCGACAGCTGAGAGC 659
Db 357 CysAlaGlyLeuGlyCysTrpGlnLeuArgGly 366

RESULT 6

024507 PRELIMINARY; PRT; 414 AA.
ID 024507
AC 024507:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Sterol carrier protein X-related thiolase.
GN SGPX OR CG17320.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 DB NCBI_TaxID=7227;
 RN
 RP
 RA
 RA Kitamura T., Kobayashi S., Okada M.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X95443; CA664718.1;
 DR Flybase: FBgn0015808; SCPX.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam: PF02803; thiolase_1.
 DR PROSITE: PS00098; THIOLEASE_1; 1.
 DR PROSITE: PS00737; THIOLEASE_2; 1.
 SQ SEQUENCE 414 AA; 44750 MW; 40CFAFE83611B7E CRC64;

Alignment Scores:

Pred. No.:	3,72e-23	Length:	414
Score:	30.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.55%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992a-21 (1-994) x Q24507 (1-414)

QY 570 AATCCTAGTGTGACTGATTTCAAGGACACCCACTAGGCGCTACAGTCTTGCTCAG 629
 |||||||

DB 337 AANPROSErGcIcIyLeuIeSerIySGIyHISPrOLEuGcIyAlaHrthIyLeuAlaGIn 356
 |||||||

QY 630 TGTGCAGAACTGTGCTGCAGCTGAGAGG 659
 |||||||

DB 357 CysAlaGlutLeucyStrpInLeuArgIy 366
 |||||||

RESULT 7
 ID Q24506 PRELIMINARY; PRT; 543 AA.

AC Q24506;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE Sterol carrier protein X.

GN SCPX OR CG17320.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

DB NCBI_TaxID=7227;
 [1]

RA
 RA SEQUENCE FROM N.A.

RA Kitamura T., Kobayashi S., Okada M.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL: X97685; CA66277.1;
 DR Flybase: FBgn0015808; SCPX.

DR InterPro: IPR003033; SCP2.

DR Pfam: PF02036; SCP2; 1.

DR Pfam: PF00108; thiolase_1.

DR PROSITE: PS00803; thiolase_C; 1.

DR PROSITE: PS00098; THIOLEASE_1; 1.

DR PROSITE: PS00737; THIOLEASE_2; 1.

RT SEQUENCE 543 AA; 58854 MW; 4B9BB43E77048181 CRC64;

Alignment Scores:

Pred. No.:	3.61e-23	Length:	543
Score:	30.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.55%	Indels:	0
DB:	5	Gaps:	0

US-09-835-992a-21 (1-994) x Q24506 (1-543)

QY 570 AATCCTAGTGTGACTGATTTCAAGGACACCCACTAGGCGCTACAGTCTTGCTCAG 629
 |||||||

DB 337 AANPROSErGcIcIyLeuIeSerIySGIyHISPrOLEuGcIyAlaHrthIyLeuAlaGIn 356
 |||||||

QY 630 TGTGCAGAACTGTGCTGCAGCTGAGAGG 659
 |||||||

DB 357 CysAlaGlutLeucyStrpInLeuArgIy 366
 |||||||

RESULT 8
 ID Q9VJ43 PRELIMINARY; PRT; 544 AA.

AC Q9VJ43;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE CG17320 protein (LD10783p).

GN SCPX OR CG17320.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

DB NCBI_TaxID=7227;
 [1]

RA
 RA SEQUENCE FROM N.A.

RA STRAIN-BERKELEY;
 MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Gelinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherter S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Y., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abblitt J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beaskey S.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottlter P.,

RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doop L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Hellman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwan C.,

RA Jatali M., Kalush F., Karpen C.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svrtskas R., Tector C., Turner R., Ventier E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Ventier J.C.;
 "The genome sequence of Drosophila melanogaster";
 Science 287:2185-2195(2000).
 [2]
 RA SEQUENCE FROM N.A.
 RA STRAIN-BERKELEY;
 Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farrant D., Frise E.,
 RA George R., Gonzalez M., Guarn H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,

RA Patcl S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Coinliker S.,
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003660; AAF53713.1; -
DR EMBL: AY089539; AAL90277.1; -
DR FlyBase: FBgn0015808; ScpX.
DR InterPro: IPR003033; SCP2.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF02036; SCP2; 1.
DR Pfam: PF02803; thiolase; 1.
DR PROSITE: PS00098; thiolase_1; 1.
DR PROSITE: PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 544 AA; 59007 MW; A14B0A7D2185E9C2 CRC64;

Alignment Scores:

Pred. No.:	3	61e-23	Length:	544
Score:	30.00	Matches:	30	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	9.55%	Indels:	0	
DB:	5	Gaps:	0	

US-09-835-992a-21 (1-994) x Q9VJ43 (1-544)

OY 570 AATCGTAGTGCTGATTTCAAGGACACCCAGCCCTACAGCTTGTCTCAG 629
|||||
DB 337 AsnProsergylglyleuileserlyglYHisProleuglyAlatnnglyleuAlaGln 356
|||||
OY 630 TGTGCAAGACTCTGCTGGCAGCTGAGAGG 659
|||||
DB 357 CysAlaGluLeuCystrpGlnLeuArgGly 366
|||||

RESULT 9

ID O033332 PRELIMINARY; PRT: 401 AA.

AC O033332;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE LIPid-transfer protein.

CN LIP1 OR RV2790C OR MY0002.55C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RX MEDLINE-98295987; PubMed-9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,

RA Davies K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Horneby T., Jorgels K., Kiroh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborn J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,

RA Saluon J.E., Taylor K., Whitehead S., Barrell B.G.;
RT Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 353:537-544(1998).
DR EMBL: AL008967; CAA15585.1; -
DR HSSP: P27796; IAFY.
DR Tuberculist: RV2790C; -
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase; 1.
DR PROSITE: PS02803; thiolase_C; 1.
DR PROSITE: PS00098; THIOLEASE_1; 1.
DR PROSITE: PS00737; THIOLEASE_2; 1.
KM Complete proteome.
SQ SEQUENCE 401 AA; 42904 MW; 6059E0B995607E CRC64;

Alignment Scores:

Pred. No.:	1	25e-16	Length:	401
Score:	24.00	Matches:	24	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	7.64%	Indels:	0	
DB:	16	Gaps:	0	

US-09-835-992a-21 (1-994) x O033332 (1-401)

OY 570 AATCGTAGTGCTGATTTCAAGGACACCCAGCCCTACAGCTTGTCTCAG 629
|||||
DB 338 AsnProsergylglyleuileserlyglYHisProleuglyAlatnnglyleuAlaGln 357
|||||
OY 630 TGTGCAAGACTCTGCTGGCAGCTGAGAGG 641
|||||
DB 358 CysAlaGluLeu 361
|||||

RESULT 10

ID O02353 PRELIMINARY; PRT: 412 AA.

AC O02353;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE 3-keto-acyl-CoA thiolase.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-N2;

RX MEDLINE-97296386; PubMed-9151950;

RA Bun-Ya M., Maebuchi M., Hashimoto T., Yokota S., Kamiryo T.;

RT "A second isoform of 3-ketoacyl-CoA thiolase found in Caenorhabditis

RT elegans, which is similar to sterol carrier protein 2.";

RT Eur. J. Biochem. 245:252-259(1997).

RN [2]

RP SEQUENCE FROM N.A.

RA Snye R.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: D86473; BAA20377.1; -

DR EMBL: AL023847; CAA19548.1; -

DR InterPro: IPR002155; Thiolase.

DR Pfam: PF00108; thiolase; 1.

DR Pfam: PF02803; thiolase_C; 1.

DR PROSITE: PS00098; THIOLEASE_1; UNKNOWN_1.

DR PROSITE: PS00737; THIOLEASE_2; 1.
SQ SEQUENCE 412 AA; 44386 MW; 6B9DA9BA18ABF446 CRC64;

Alignment Scores:

Pred. No.:	1	85e-14	Length:	412
Score:	22.00	Matches:	22	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	7.01%	Indels:	0	
DB:	5	Gaps:	0	

US-09-835-992a-21 (1-994) x O02353 (1-412)

OY 540 GATATACATATGAGAGAAAGTGGCTCAATCTAGTGGACGATTTCAAGGCA 599
|||||
DB 332 AspnAnthrlyrglyglylystrpValIleAnProsergylglyleuileserlyGly 351
|||||
OY 600 CACCCA 605
|||||
DB 352 HisPro 353
|||||

RESULT 11

ID O8VJDS PRELIMINARY; PRT: 400 AA.

AC O8VJDS;

DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Nonspecific lipid-transfer protein.
 GN MT2860.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE007112; AAK4719.1; -.
 DR TIGR; MT2860; -.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam; PF02803; thiolase_C; 1.
 DR PROSITE: PS00098; THIOLEASE_1; UNKNOWN_1.
 DR PROSITE: PS00737; THIOLEASE_2; UNKNOWN_1.
 SQ SEQUENCE 400 AA; 42833 MW; 8A75C1BB204DF7A2 CRC64;

Alignment Scores:

Pred. No.: 2,27e-13 Length: 400
 Score: 21.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.69% Indels: 0
 DB: Gaps: 0

US-09-835-992a-21 (1-994) x Q8VUD5 (1-400)

QY 570 AATCGTAGTGCTGACTGATTCAAGGACACCCACTAGCGCTCAGTCTGCTCAG 629
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 338 AANPROSERGLYGLYLEULEISERLYSGLYHISPROLEUGLYALATHRGLYLEUALAGLN 357

QY 630 TGT 632

DB 358 CYS 358

RESULT 12

Q9EZ05 PRELIMINARY; PRT; 132 AA.

AC 09EZ05;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 13.3 kDa protein (Fragment).
 OS Streptomyces aureofaciens.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CM3239;
 RA Kormanec J., Novakova R., Homerova D., Rezuchova B.;
 RT "Putative meta-cleavage compound hydrolase gene is dependent on a
 RT sporulation specific sigma RpoZ in Streptomyces aureofaciens.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF306830; AAG42771.1; -.
 DR HSSP; P27796; IAFY.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam; PF02803; thiolase_C; 1.
 DR PROSITE: PS00737; THIOLEASE_2; 1.
 KW Hypothetical protein.
 FT NON_TER 1

SQ SEQUENCE 132 AA; 13296 MW; E1173B31584DE487 CRC64;

Alignment Scores:

Pred. No.: 3.17e-12 Length: 132
 Score: 20.00 Matches: 20
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.37% Indels: 0
 DB: Gaps: 0

US-09-835-992a-21 (1-994) x Q9EZ05 (1-132)

QY 570 AATCGTAGTGCTGACTGATTCAAGGACACCCACTAGCGCTCAGTCTGCTCAG 629
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 72 AANPROSERGLYGLYLEULEISERLYSGLYHISPROLEUGLYALATHRGLYLEUALAGLN 91

RESULT 13

Q29224 PRELIMINARY; PRT; 152 AA.

AC 029224;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Sterol carrier protein-X/sterol carrier protein-2 (Fragment).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Cetartiodactyla; Suidae; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINE;
 RC MEDLINE=96327607; PubMed=8672129;
 RA Wintencoe A.K., Fredholm M., Davies W.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library";
 RL Mamm. Genome 7:509-517(1996).
 DR EMBL; F14816; CAA23274.1; -.
 DR InterPro: IPR002155; Thiolase.
 DR Pfam; PF02803; thiolase_C; 1.
 FT NON_TER 1
 FT NON_TER 152
 SQ SEQUENCE 152 AA; 16276 MW; 9632305FE9FCA6A CRC64;

Alignment Scores:

Pred. No.: 3.81e-11 Length: 152
 Score: 19.00 Matches: 19
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.05% Indels: 0
 DB: Gaps: 0

US-09-835-992a-21 (1-994) x Q29224 (1-152)

QY 363 GTTGGCTTGATGATGATAAGAGCTGCAGAAATGCTATGAGAAATCTGGCTG 419
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 72 VALGLYPHEASMERSELYSGLYHISPROLEUGLYALATHRGLYLEUALAGLN 90

RESULT 14

Q97VJ9 PRELIMINARY; PRT; 387 AA.

AC 097VJ9;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Acetyl-CoA C-acetyltransferase (acetoacetyl-CoA thiolase)
 DE (acac-7) (EC 2.3.1.9).
 GN ACAB-7 OR SSO2625.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moers A., Frausto G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.T., Medina N., Peng X.,
RA Thir-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT *The complete genome of the crenarchaeon Sulfolobus solfataricus P2.*;
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL: AE006858; AKA2745.1.;
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; Thiolase; 1.
DR Pfam: PF02803; Thiolase; 1.
KW transferase; Acyltransferase; Complete proteome.
SQ SEQUENCE 387 AA; 41892 MW; 4F6E505FEDB3522F CRC64;

Alignment Scores:

Pred. No.:	0.207	Length:	387
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.18%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992a-21 (1-994) x Q97VJ9 (1-387)

QY 594 AAGGACACCCACTAGGCGCTACAGGTCTT 623

DB 333 LysGlyHisProLeuGlyAlaIrnGlyLeu 342

RESULT 15

Q96XU9 PRELIMINARY; PRT; 387 AA.
ID Q96XU9;
AC Q96XU9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative nonspecific lipid-transfer protein.
GN ST2418.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekino M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagi M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT *Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.*;
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000990; BAB67528.1.;
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; Thiolase; 1.
DR Pfam: PF02803; Thiolase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 387 AA; 42071 MW; DF3372F987C42D69 CRC64;

Alignment Scores:

Pred. No.:	0.207	Length:	387
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.18%	Indels:	0
DB:	17	Gaps:	0

US-09-835-992a-21 (1-994) x Q96XU9 (1-387)

QY 594 AAGGACACCCACTAGGCGCTACAGGTCTT 623

DB 333 LysGlyHisProLeuGlyAlaIrnGlyLeu 342

Search completed: January 14, 2003, 17:37:03
Job time: 77.9914 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:24 : Search time 41.5252 Seconds
(without alignments)
5410.226 Million cell updates/sec

Title: US-09-835-992a-22

Perfect score: 180
Sequence: 1 ggcacaaaataattattcna.....taaatcgacaataaaga 843

Scoring table:
OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+np.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US0983592/runat.14012003.161525.3979/app_query_dev_us-09-835-992a-22.olig.rag -MINMATCH=0.1 -LOOPT=0
-DB=A.GeneSeq.101002 -OFMT=fasta -SUFFIX=olig.rag -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0983592.ecgn.1.1.121.etruncat.14012003.161525.3979 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT-120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A.GeneSeq.101002.*
1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	7	3.8	56	20	AA02476	Clone selected aft
2	7	3.9	59	22	ABP03819	Human ORFX protein
3	7	3.9	172	22	AA052473	Mycobacterium tube
C 4	7	3.8	185	21	AAV74955	Neisseria meningit
5	7	3.9	188	20	AAW54491	M. tuberculosis RP
6	7	3.9	584	22	ABG05212	Novel human diagno
7	7	3.9	594	22	ABG21582	Novel human diagno
8	7	3.9	691	21	AB07698	A Neisseria mening
9	7	3.9	722	21	AB07697	A Neisseria mening
10	7	3.9	1718	22	ABG24524	Novel human diagno
11	6	3.3	15	21	AA033322	Antimicrobial pep
12	6	3.3	25	21	AA069422	Antimicrobial pep
13	6	3.3	25	21	AA069423	Antimicrobial pep
14	6	3.3	25	21	AA069424	Antimicrobial pep
15	6	3.3	25	21	AA069425	Antimicrobial pep
16	6	3.3	35	16	AA074455	Parathyroid hormon
17	6	3.3	35	22	ABG01026	Novel human diagno
18	6	3.3	36	23	ABP30830	Streptococcus poly
C 19	6	3.3	39	21	AA038159	Human secreted pro
20	6	3.3	41	21	AA044645	Human secreted pro
21	6	3.3	42	21	AA012320	Human secreted pro
22	6	3.3	43	23	AA021238	Human gene 1 encod
C 23	6	3.3	46	22	AA090896	Human immune/hema
C 24	6	3.3	46	23	ABG46482	Human peptide enco
25	6	3.3	50	23	ABP04160	Human ORFX protein
C 26	6	3.3	57	22	ABG12659	Human polypeptide
C 27	6	3.3	58	22	AA004926	Human polypeptide
28	6	3.3	61	22	AB014799	Human nervous syst
29	6	3.3	64	22	AA003024	Human polypeptide
C 30	6	3.3	65	22	AA051896	Propionibacterium
C 31	6	3.3	65	23	AA080965	Toxoplasma graz an
C 32	6	3.3	67	23	ABP01643	Human ORFX protein
C 33	6	3.3	68	23	ABP04033	Human ORFX protein
C 34	6	3.3	72	23	ABG60075	Human DTRP polype
35	6	3.3	73	21	AA040726	Human ORFX ORF490
C 36	6	3.3	74	22	AA086612	Novel human connec
C 37	6	3.3	74	22	AA000202	Human polypeptide
C 38	6	3.3	78	22	AA003840	Human polypeptide
C 39	6	3.3	78	23	ABP03178	Human ORFX protein
40	6	3.3	79	23	ABP42461	Human ovarian anti
C 41	6	3.3	85	23	ABP33738	Human ORF2711 prot
C 42	6	3.3	86	20	AA079853	Human secreted pro
C 43	6	3.3	86	20	AA079852	Human secreted pro
C 44	6	3.3	86	22	AA083146	Human immune/hema
C 45	6	3.3	87	22	AA003194	Human polypeptide

ALIGNMENTS

RESULT 1
AA02476
ID AA02476 standard; Protein; 56 AA.

XX AA02476;

XX 15-JUL-1999 (first entry)

XX Clone selected after panning a NK library of the invention.

XX Screening; functional polypeptide; ligand; non-functional;

XX enrichment; single chain antibody; Scrv.

XX Synthetic.

XX MO9920749-A1.

XX 29-APR-1999.

XX

PF 20-OCT-1998; 98WO-GB03135.
XX
PR 21-NOV-1997; 97US-0066729.
PR 20-OCT-1997; 97GB-0022131.
PR 13-NOV-1997; 97US-0065428.
XX
PA (MED-) MEDICAL RES COUNCIL.
XX
PI Tomlinson I, Winter G;
XX
DR WPI; 1999-288302/24.
XX
PT Screening for functional polypeptides which bind a ligand
XX
PS Example 3; Fig 4; 67pp; English.
XX
CC The specification describes a method for screening for functional
CC polypeptides which bind a ligand. The method comprises contacting a
CC repertoire of polypeptides with a generic ligand, and then screening
CC selected functional polypeptides with a target ligand. The method
CC permits the removal from a chosen repertoire of polypeptides, those
CC which are non-functional, e.g. as a result of the introduction of
CC frame-shift mutations, stop codons, folding mutants or expression
CC mutants which would be or are incapable of binding to any target
CC ligand. The method also permits the enrichment of a chosen repertoire
CC of polypeptides for those polypeptides which are functional, well folded
CC and highly expressed. The polypeptides obtained can be used in
CC diagnostic, prophylactic and therapeutic procedures. AAY02473-Y02525
CC represent clones selected after panning primary and somatic NMK
CC libraries with 5 antigens (bovine ubiquitin, rat BIP, bovine histone,
CC NIP-BSA and hen egg lysozyme).
XX
SQ Sequence 56 AA:

Alignment Scores:
Pred. No.: 49.5 Length: 56
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: Gaps: 0

US-09-835-992a-22 (1-843) x AAY02476 (1-56)
QY 802 AAAGNGGGGAGGTTGTTT 782
DB 21 LysglyglyglyArgLeuPhe 27

RESULT 2
ABP05819
ID ABP05819 standard; Protein: 59 AA.
XX
AC ABP05819;
XX
DT 24-JUN-2002 (first entry)
DE Human ORFX protein sequence SEQ ID NO:11620.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KM hypertension; hypothyroidism; cholesterol ester storage disease;
KM immune deficiency; immune disorder; infectious disease;
KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KM myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX

PF 29-MAY-2001; 2001WO-US10836.
XX
PR 30-MAY-2000; 2000US-206132P.
PR 29-AUG-2000; 2000US-228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach MD;
XX
DR WPI; 2002-106308/14.
XX
DR N-PSDB; ABN21571.
XX
PT Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders
XX
PS Disclosure; SEQ ID 11620; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where x is 1-11491 (see Table 1
CC in the specification)). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 59 AA:

Alignment Scores:
Pred. No.: 49.1 Length: 59
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: Gaps: 0

US-09-835-992a-22 (1-843) x ABP05819 (1-59)
QY 773 CCCCTAAAAACAACCTCC 793
DB 34 ProleuylslysglnProser 40

RESULT 3
AAB52473
ID AAB52473 standard; protein: 172 AA.
XX
AC AAB52473;
XX
DT 23-FEB-2001 (first entry)
DE Mycobacterium tuberculosis secreted protein #38.
XX
KM Mycobacterium tuberculosis secreted protein; MTSF; vaccine.
OS Mycobacterium tuberculosis.

```

XX PN WO20006143-A1.
XX PD 09-NOV-2000.
XX PF 04-MAY-2000; 2000WO-US12197.
XX PR 04-MAY-1999; 9905-0132479.
XX PR 04-MAY-1999; 9905-0132503.
XX PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
XX PI Gennaro ML, Gomez MJ;
XX DR WPI; 2001-007151/01.
XX PT Novel Mycobacterium tuberculosis secreted polypeptides and
XX PT polynucleotides useful in diagnosis, treatment and prophylaxis of
XX PT tuberculosis.
XX PS Claim 1; Fig 1; 60pp; English.
XX CC The present invention relates to Mycobacterium tuberculosis secreted
XX CC proteins (MTPs), where the polypeptide has M. tuberculosis specific
XX CC antigenic and immunogenic properties. Compositions of the invention may
XX CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
XX CC vaccine against M. tuberculosis infection.
XX SQ Sequence 172 AA:

Alignment Scores:
Pred. No.: 41.7 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 22 Gaps: 0

US-09-835-992a-22 (1-843) x AAB52473 (1-172)
OY 786 AACCTTCCGCCNCCTTTGGCC 806
DB 73 AenLeuProProLeuAa 79

RESULT 4
AAW74955
ID AAW74955 standard; Protein; 185 AA.
XX AC AAW74955;
XX DT 21-MAR-2000 (first entry)
XX DE Neisseria meningitidis ORF 507 protein sequence SEQ ID NO:1384.
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
XX KW antibacterial; gene therapy.
XX OS Neisseria meningitidis.
XX PN WO9957280-A2.
XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US09346.
XX PR 01-MAY-1998; 98US-0083758.
XX PR 31-JUL-1998; 98US-0094869.
XX PR 02-SEP-1998; 98US-0098994.
XX PR 02-SEP-1998; 98US-0099062.
XX PR 09-OCT-1998; 98US-0103749.
XX PR 09-OCT-1998; 98US-0103794.
XX PR 09-OCT-1998; 98US-0103796.
XX PA

```

```

PR 25-FEB-1999; 99US-0121528.
XX PA (CHIR) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Maignani V, Mora M;
XX PI Petersen J, Pizzo M, Rappoli R, Ratti G, Scalato E, Scarselli M;
XX PI Tettelin H, Venter JC;
XX DR WPI; 2000-062150/05.
XX DR N-PSDB; AA53717.
XX PT Novel Neisserial polypeptides predicted to be useful antigens for
XX PT vaccines and diagnostics.
XX PS Claim 2; Page 752-753; 1453pp; English.
XX CC AA253015 to AA254536, AA254577 to AA254615, and AAW74253 to AAW75941
XX CC represent novel Neisseria meningitidis and N. gonorrhoeae polynucleotides
XX CC and polypeptides. AA254537 to AA254576 and AA254616 to AA25473 represent
XX CC PCR primers used in the exemplification of the present invention. The
XX CC polypeptides, the polynucleotides, antibodies and compositions of
XX CC the invention can be used as vaccines, as diagnostic reagents, and as
XX CC immunogenic compositions. The polypeptides can be used in the
XX CC manufacture of medicaments for treating or preventing infection due to
XX CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
XX CC presence of Neisseria bacteria, or to raise antibodies. They may also
XX CC be used to screen for agonists or antagonists, which may themselves
XX CC have use as antibacterial agents. The polynucleotides of the invention
XX CC may also be used in gene therapy protocols.
XX SQ Sequence 185 AA:

Alignment Scores:
Pred. No.: 41.2 Length: 185
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-22 (1-843) x AAW74955 (1-185)
OY 789 GGTGTGTTTTTAAGGGGGGA 769
DB 10 GlyCyspheLeuArgGlyGly 16

RESULT 5
AAW95491
ID AAW95491 standard; Protein; 188 AA.
XX AC AAW95491;
XX DT 26-MAR-1999 (first entry)
XX DE M. tuberculosis RP-factor-like gene product (MtubM7V008).
XX KW M. tuberculosis RP-factor; latent cell; growth-stimulation; receptor;
XX KW RP-factor; resuscitate; latent cell; growth-stimulation; receptor;
XX KW convertase; vaccine; immunotherapy; prophylaxis; excitant; antibiotic;
XX KW antimicrobial; tuberculosis; food; pharmaceutical; culture additive.
XX OS Mycobacterium tuberculosis.
XX PN WO9855624-A1.
XX PD 10-DEC-1998.
XX PF 03-JUN-1998; 98WO-GB01619.
XX PR 27-MAY-1998; 98GB-0011221.
XX PR 04-JUN-1997; 97GB-0011389.
XX PA (UYWA-) UNIV WALES.

```

```

XX Kaprelyants AS, Kell DB, Mukamolova GV, Young DI;
PI Young M;
XX
XX WPI: 1999-070270/06.
XX
XX New bacterial resuscitation factors - useful for stimulating latent
PT bacteria and growth, particularly for use as antimicrobials; also
PT vaccines comprising bacteria with mutations in resuscitation factor
PT genes
XX
XX Disclosure; Fig 1A; 76pp; English.
XX
XX The invention relates to RP-factors which are substances that can
CC resuscitate dormant, moribund or latent cells, possibly also having
CC growth-stimulating activity. Host cells containing a vector comprising
CC the RP-factor or its receptor encoding nucleic acid can be used for the
CC recombinant production of the RP-factor. RP-factors, their receptors or
CC convertases, antibodies (against the RP-factors and RP-factor receptors
CC or convertases), antagonists and agonists, are useful in vaccines and for
CC immunotherapy, diagnosis and prophylaxis. They are also useful as
CC excipients, generally as antimicrobials and especially for infections
CC associated with latency. They can be used as potentiators of antibiotics
CC such as isoniazid, streptomycin etc., in treatment of tuberculosis.
CC RP-factor can also be used: to determine the microbiological quality of
CC foods, pharmaceuticals, medical products; as culture additive for
CC bacteria; to stimulate growth and/or to resuscitate microorganisms and to
CC produce libraries of biomolecules and microorganisms (which may then be
CC screened for useful products). Probes complementary to the RP-factor
CC nucleic acid are used identify and clone other RP-factor genes. The
CC inventions may also be used to prevent bacterial resuscitation. Breaking
CC dormancy with RP-factor facilitates detection, culture and enumeration of
CC many bacteria. The present sequence represents a predicted amino acid
CC sequence of a RP-factor-like gene product from M. tuberculosis.
XX
XX Sequence 188 AA:
SQ
XX
XX Alignment Scores:
XX Pred. No.: 41.1 Length: 188
XX Score: 7.00 Matches: 7
XX Percent Similarity: 100.00% Conservativity: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3 89% Indels: 0
XX DB: 20 Gaps: 0
XX
XX US-09-835-992A-22 (1-843) x AAW95491 (1-188)
OY 786 AACCTTCCGCCNCGTTGGCC 806
DB 89 AasnleuproProleuNa1a 95
XX
XX RESULT 6
XX ABG05212
XX ID ABG05212 standard; Protein; 584 AA.
XX
XX ABG05212;
XX
XX 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #5203.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX

```

PR	23-AUG-2000; 2000US-0649167.		
XX	(HYSE-) HYSEQ INC.		
XX			
XX	Drmanac RT, Liu C, Tang YT;		
XX			
XX	WPI: 2001-639362/73.		
DR	N-PFDB; AAS69399.		
XX			
PT	New isolated polynucleotide and encoded polypeptides, useful in		
PT	diagnostics, forensics, gene mapping, identification of mutations		
PT	responsible for genetic disorders or other traits and to assess		
PT	biodiversity -		
XX			
PS	Claim 20; SEQ ID No 35571; 103bp; English.		
XX			
CC	The invention relates to isolated polynucleotide (I) and		
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,		
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome		
CC	and gene mapping, and in recombinant production of (II). The		
CC	polynucleotides are also used in diagnostics as expressed sequence tags		
CC	for identifying expressed genes. (I) is useful in gene therapy techniques		
CC	to restore normal activity of (II) or to treat disease states involving		
CC	(II). (II) is useful for generating antibodies against it, detecting or		
CC	quantitating a polypeptide in tissue, as molecular weight markers and as		
CC	a food supplement. (II) and its binding partners are useful in medical		
CC	imaging of sites expressing (II). (I) and (II) are useful for treating		
CC	disorders involving aberrant protein expression or biological activity.		
CC	The polypeptide and polynucleotide sequences have applications in		
CC	diagnostics, forensics, gene mapping, identification of mutations		
CC	responsible for genetic disorders or other traits to assess biodiversity		
CC	and to produce other types of data and products dependent on DNA and		
CC	amino acid sequences. ABG0010-ABG30377 represent novel human		
CC	diagnostic amino acid sequences of the invention.		
CC	Note: The sequence data for this patent did not appear in the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences .		
XX			
XX	Sequence 584 AA:		
	Alignment Scores:		
	Pred. No.:	34.6	Length: 584
	Score:	7.00	Matches: 7
	Percent Similarity:	100.00%	Conservative: 0
	Best Local Similarity:	100.00%	Mismatches: 0
	Query Match:	3.89%	Indels: 0
	DB:	22	Gaps: 0
	US-09-835-992A-22 (1-843) x ABG05212 (1-584)		
OY	670 TTACCATTTTACCTTGTAAG 690		
DB	172 LeuProPheTyrIleucllyLys 178		
	RESULT 7		
	ABG21682		
ID	ABG21682 standard; Protein; 594 AA.		
XX			
AC	ABG21682;		
XX			
XX	18-FEB-2002 (first entry)		
XX			
DE			
XX	Novel human diagnostic protein #21673.		
XX			
KM	Human; chromosome mapping; gene mapping; gene therapy; forensic;		
KW	food supplement; medical imaging; diagnostic; genetic disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200175067-A2.		
XX			
PD	11-OCT-2001.		
XX			

PF	30-MAR-2001; 2001WO-US08631.	
PR	31-MAR-2000; 2000US-0540217.	
XX	23-AUG-2000; 2000US-0649167.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Drmnmac RT, Liu C, Tang YT;	
DR	WPI; 2001-639362/73.	
DR	N-PSDB; AAS85869.	
XX		
PT	New isolated polynucleotide and encoded polypeptides, useful in	
PT	diagnostics, forensics, gene mapping, identification of mutations	
PT	responsible for genetic disorders or other traits and to assess	
PT	biodiversity -	
XX		
PS	Claim 20; SEQ ID NO 52041; 103pp; English.	
XX		
CC	The invention relates to isolated polynucleotide (I) and	
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,	
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome	
CC	and gene mapping, and in recombinant production of (II). The	
CC	polynucleotides are also used in diagnostics as expressed sequence tags	
CC	for identifying expressed genes. (I) is useful in gene therapy techniques	
CC	to restore normal activity of (II) or to treat disease states involving	
CC	quantitating a polypeptide in tissue, as molecular weight markers and as	
CC	a food supplement. (II) and its binding partners are useful in medical	
CC	imaging of sites expressing (II). (I) and (II) are useful for treating	
CC	disorders involving aberrant protein expression or biological activity.	
CC	The polypeptide and polynucleotide sequences have applications in	
CC	diagnostics, forensics, gene mapping, identification of mutations	
CC	responsible for genetic disorders or other traits to assess biodiversity	
CC	and to produce other types of data and products dependent on DNA and	
CC	amino acid sequences. ABG00010-ABG30377 represent novel human	
CC	diagnostic amino acid sequences of the invention.	
CC	Note: The sequence data for this patent did not appear in the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		
SO	Sequence 594 AA;	
XX		
Alignment Scores:		
Pred. NO.:	34.5	Length: 594
Score:	7.00	Matches: 7
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	3.89%	Indels: 0
DB:	22	Gaps: 0
XX		
US-09-835-992A-22 (1-843) x ABG21682 (1-594)		
Oy	670 TTACCATTTTACCTTGGTAAC 650	
Db	TTTTTTTTTTTTTTTT	
	172 LeuProPheTyrlcuGlyLys 178	
XX		
RESULT 8		
AAB07698		
ID	AAB07698 standard; Protein: 691 AA.	
XX		
AC	AAB07698;	
XX		
DT	07-NOV-2000 (first entry)	
XX		
A	Neisseria meningitidis BAS053 polypeptide.	
XX		
XX	BASB053; Neisseria meningitidis infection; vaccine.	
XX		
OS	Neisseria meningitidis.	
XX		
PN	WO200042193-A1.	

```

PD 20-JUL-2000.
PF 10-JAN-2000; 2000WO-EP00137.
PR 15-JAN-1999; 99GB-0000959.
PR 28-JAN-1999; 99GB-0001903.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA
XX
XX Ruelle J;
PI
XX
XX MPI: 2000-476062/41.
DR N-PSDB: AAAS9217.
XX
XX New Neisseria meningitidis polypeptide useful for diagnosis of
PT Neisseria infection and for development of vaccines against such
PT infection -
XX
XX Claim 3; Page 56; 92pp; English.
XX
XX The present sequence represents a Neisseria meningitidis BASB053
CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
CC for BASB053 may be identified in a biological sample in order to
CC diagnose a Neisseria meningitidis infection in an animal. The BASB053
CC polypeptides and polynucleotides may be used as vaccines, for
CC generating an immune response in an animal. A composition comprising
CC at least one antibody immunospecific for BASB053 may be used to
CC treat humans infected with Neisseria meningitidis.
XX
XX Sequence 691 AA;
SO
Alignment Scores:
Pred. No.: 33..7 Length: 691
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatvie: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: Gaps: 0
US-09-835-992A-22 (1-843) x AAB07698 (1-691)
QY 823 CTTAATCGGACATTTAAAG 843
DB 535 LeuAsnProAsnAsnLysLys 541
RESULT 9
AAB07697
ID AAB07697 standard; Protein: 722 AA.
XX
XX AAB07697;
AC
XX
XX 07-NOV-2000 (first entry)
DT
XX
XX A Neisseria meningitidis BASB053 polypeptide.
DE
XX
XX BASB053; Neisseria meningitidis infection; vaccine.
KW
XX
XX Neisseria meningitidis.
OS
XX
XX WO200042193-A1.
PN
XX
XX 20-JUL-2000.
PD
XX
XX 10-JAN-2000; 2000WO-EP00137.
PF
XX
XX 15-JAN-1999; 99GB-0000959.
PR
XX
XX 28-JAN-1999; 99GB-0001903.
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PI
XX
XX Ruelle J;
DR
MPI: 2000-476062/41.

```

DR N-PSDB; AAS59216.
XX
PT New *Neisseria meningitidis* polypeptide useful for diagnosis of
PT *Neisseria* infection and for development of vaccines against such
PT infection -
XX
PS Claim 3; Page 55-56; 92pp; English.
XX
CC The present sequence represents a *Neisseria meningitidis* BASB053
CC polypeptide. The BASB053 polypeptide, or an antibody immunospecific
CC for BASB053 may be identified in a biological sample in order to
CC diagnose a *Neisseria meningitidis* infection in an animal. The BASB053
CC polypeptides and polynucleotides may be used as vaccines, for
CC generating an immune response in an animal. A composition comprising
CC at least one antibody immunospecific for BASB053 may be used to
CC treat humans infected with *Neisseria meningitidis*.
SQ Sequence 722 AA;

Alignment Scores:
Pred. No.: 33.5 Length: 722
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-22 (1-843) x AAB07697 (1-722)
OY 823 CTAATCCGACATATAAAA 843
ID 566 LeuAsnProAsnAsnLys 572
ABG24524
ID ABG24524 standard; Protein; 1718 AA.
XX
AC ABG24524;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24515.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88711.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 54883; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (II) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 1718 AA;

Alignment Scores:
Pred. No.: 29.4 Length: 1718
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 22 Gaps: 0

US-09-835-992a-22 (1-843) x ABG24524 (1-1718)
OY 670 TTACCATTTTACCTTGTAAG 690
ID 220 LeuProPheTyrLeuGlyLys 226
ABG03322
ID AAB03322 standard; Peptide; 15 AA.
XX
AC AAB03322;
XX
DT 08-SEP-2000 (first entry)
XX
DE Human epitope Rh50 GP # 32 corresponding to residues 311-325.
XX
KW Human; rhesus blood group system; Rh; RHD; RhCE; sickle cell disease;
KW thalassemia; Rhc; Rbc; Rbc; Rbc; alloimmunisation prevention;
KW autoimmune Rh haemolytic disease; rhesus protein; immunosuppressive;
XX
OS Homo sapiens.
XX
PN WO200032632-A2.
PD 08-JUN-2000.
XX
PF 01-DEC-1999; 99WO-GB04027.
XX
PR 01-DEC-1998; 98GB-0026378.
XX
PA (UYAB-) UNIV ABERDEEN.
PA (COMM-) COMMON SERVICES AGENCY SCOTTISH HEALTH S.
XX
PI Urbanlak SJ, Barker RN;
XX
DR WPI; 2000-412291/35.
XX
PT Composition for prevention of alloimmunization or immunosuppression of
PT a response elicited by alloimmunization or an autoimmune hemolytic
PT disease, comprises an epitope of a rhesus protein
XX
PS Disclosure; Page 78; 92pp; English.

XX Human blood contains the rhesus (Rh) blood group system, and humans can
CC either be RhD positive or negative. This can lead to complications
CC during transfusions or pregnancy if RhD negative individuals are exposed
CC to RhD positive blood, leading to them becoming immunised to produce
CC anti-D. The present invention relates to new human allo- and
CC auto-reactive T-cell epitopes (AAV99760-199769 and AAB03201-803337) from
CC RhD, Rhc, Rnc, Rhd and Rhe proteins. These epitopes bind to T-cells to
CC elicit an immune response, i.e. immunisation. These epitopes can be used
CC as a vaccine for the prevention of alloimmunisation or immunosuppression
CC of a response elicited by alloimmunisation or an autoimmune haemolytic
CC disease. Examples of autoimmune haemolytic diseases are sickle cell
CC disease and thalassemia.

XX
SQ Sequence 15 AA:

Alignment Scores:
Pred. No.: 804 Length: 15
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-22 (1-843) x AAB03322 (1-15)

OY 412 AATTTTACCCCTA 429
DB 2 LysPheLeuThrProLeu 7
|||||

RESULT 12
AAV69422
ID AAV69422 standard; peptide: 25 AA.
XX
AC AAV69422:
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1.
XX
KW Antimicrobial: Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; Temporin-B; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus.
XX
OS Rana catesbeiana.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 1; Page 35; 61pp; English.
XX
CC AAV69422-75 represent antibacterial peptides (termed Rana peptides)

CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or
CC Temporin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit Escherichia coli, Staphylococcus aureus, Candida albicans,
CC Enterococcus sp., Bacteroides fragilis, Pseudomonas aeruginosa,
CC Klebsiella pneumoniae, Serratia marcescens, Mycobacterium tuberculosis,
CC Streptococcus pneumoniae, Streptococcus pyogenes, Haemophilus
CC influenzae, and Staphylococcus saprophyticus.

XX
SQ Sequence 25 AA:

Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0

US-09-835-992a-22 (1-843) x AAV69422 (1-25)

OY 96 TTAATAAACCTNGCAAA 113
DB 6 LeuLysAsnLeuGlyLys 11
|||||

RESULT 13
AAV69423
ID AAV69423 standard; peptide: 25 AA.
XX
AC AAV69423:
XX
DT 03-JUL-2000 (first entry)
XX
DE Antimicrobial peptide derived from the protein Ranatuerin-1G.
XX
KW Antimicrobial: Rana peptide; Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Temporin-A; Temporin-B; American frog;
KW bacterial infection; Escherichia coli; Staphylococcus aureus;
KW Candida albicans; Enterococcus; Bacteroides fragilis;
KW Pseudomonas aeruginosa; Klebsiella pneumoniae; Serratia marcescens;
KW Mycobacterium tuberculosis; Streptococcus pneumoniae;
KW Streptococcus pyogenes; Haemophilus influenzae;
KW Staphylococcus saprophyticus; Ranatuerin-1G.
XX
OS Rana gryllo.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
PF 13-AUG-1999; 99WO-US18575.
XX
PR 14-AUG-1998; 98US-0096607.
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
XX
PI Conlon JM;
XX
DR WPI; 2000-237449/20.
XX
PT Antibacterial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit E. coli, S. aureus and C. albicans
PT infection -
XX
PS Claim 1; Page 35; 61pp; English.
XX
CC AAV69422-75 represent antibacterial peptides (termed Rana peptides)
CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or

CC Tempurin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used to
CC to inhibit *Escherichia coli*, *Staphylococcus aureus*, *Candida albicans*,
CC *Enterococcus sp.*, *Bacteroides fragilis*, *Pseudomonas aeruginosa*,
CC *Klebsiella pneumoniae*, *Serratia marcescens*, *Mycobacterium tuberculosis*,
CC *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Haemophilus*
CC *Influenzae*, and *Staphylococcus saprophyticus*.
SQ Sequence 25 AA:
Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0
US-09-835-992a-22 (1-843) x AAY69423 (1-25)
QY 96 TTAACCTGAGAAA 113
DB 6 LeuLysAsnLeuGlyLys 11
RESULT 14
ID AAY69424 standard; peptide: 25 AA.
AC AAY69424;
XX
DT 03-JUL-2000 (first entry)
DE Antimicrobial peptide derived from the protein Ranatuerin-1C.
XX
XX Antibacterial: Rana peptide: Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Tempurin-A; Tempurin-B; American frog;
KW bacterial infection; *Escherichia coli*; *Staphylococcus aureus*;
KW *Candida albicans*; *Enterococcus*; *Bacteroides fragilis*;
KW *Pseudomonas aeruginosa*; *Klebsiella pneumoniae*; *Serratia marcescens*;
KW *Mycobacterium tuberculosis*; *Streptococcus pneumoniae*;
KW *Streptococcus pyogenes*; *Haemophilus influenzae*;
KW *Staphylococcus saprophyticus*; Ranatuerin-1C.
XX
OS Rana clamitans.
XX
PN WO200009553-A2.
XX
PD 24-FEB-2000.
XX
XX 13-AUG-1999; 99WO-US18575.
PF
XX 14-AUG-1998; 98US-0096607.
PR
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
PI
PI Conlon JM;
PI
DR WPI: 2000-237449/20.
XX
XX Antimicrobial peptides selected from 7 different Rana peptide families
FT e.g. Ranatuerin-1, used to inhibit *E. coli*, *S. aureus* and *C. albicans*
PT infection
PT
XX
PS Claim 1; Page 35; 61pp; English.
XX
CC AAY69422-75 represent antibacterial peptides (termed Rana peptides)
CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-A, or
CC Tempurin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are

CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit *Escherichia coli*, *Staphylococcus aureus*, *Candida albicans*,
CC *Enterococcus sp.*, *Bacteroides fragilis*, *Pseudomonas aeruginosa*,
CC *Klebsiella pneumoniae*, *Serratia marcescens*, *Mycobacterium tuberculosis*,
CC *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Haemophilus*
CC *Influenzae*, and *Staphylococcus saprophyticus*.
SQ Sequence 25 AA:
Alignment Scores:
Pred. No.: 744 Length: 25
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 21 Gaps: 0
US-09-835-992a-22 (1-843) x AAY69424 (1-25)
QY 96 TTAACCTGAGAAA 113
DB 6 LeuLysAsnLeuGlyLys 11
RESULT 15
ID AAY69469 standard; peptide: 25 AA.
AC AAY69469;
XX
DT 03-JUL-2000 (first entry)
DE Antimicrobial peptide derived from the protein Ranatuerin-1.
XX
XX Antibacterial: Rana peptide: Ranatuerin-1; Ranatuerin-2; Esculentin-2;
KW Brevinin-1; Ranalexin-1; Tempurin-A; Tempurin-B; American frog;
KW bacterial infection; *Escherichia coli*; *Staphylococcus aureus*;
KW *Candida albicans*; *Enterococcus*; *Bacteroides fragilis*;
KW *Pseudomonas aeruginosa*; *Klebsiella pneumoniae*; *Serratia marcescens*;
KW *Mycobacterium tuberculosis*; *Streptococcus pneumoniae*;
KW *Streptococcus pyogenes*; *Haemophilus influenzae*;
KW *Staphylococcus saprophyticus*.
XX
OS Rana catesbeiana.
XX
FH
XX Key Location/Qualifiers
FT Misc-difference 3 /note= "not specified"
FT Misc-difference 16 /note= "not specified"
FT Misc-difference 21 /note= "not specified"
FT
XX
PN WO200009553-A2.
XX
XX 24-FEB-2000.
PD
XX 13-AUG-1999; 99WO-US18575.
PF
XX 14-AUG-1998; 98US-0096607.
PR
XX
PA (BION-) BIONEERASKA INC.
PA (CONL/) CONLON J M.
PI
PI Conlon JM;
PI
DR WPI: 2000-237449/20.
XX
XX Antimicrobial peptides selected from 7 different Rana peptide families
PT e.g. Ranatuerin-1, used to inhibit *E. coli*, *S. aureus* and *C. albicans*
PT infection
PT
XX
PS Claim 2; Page 37; 61pp; English.

XX AAY69422-75 represent antibacterial peptides (termed Rana peptides)
CC which are derived from the Rana peptide families Ranatuerin-1,
CC Ranatuerin-2, Esculentin-2, Brevinin-1, Ranalexin-1, Temporin-A, or
CC Temporin-B. Each peptide is in the form of a C-terminus carboxylic
CC acid, a C-terminus amide or a C-terminus ester. The peptides are
CC derived from the skin of American frogs. The peptides can be used to
CC treat a bacterial infection in a patient. The peptides can be used
CC to inhibit *Escherichia coli*, *Staphylococcus aureus*, *Candida albicans*,
CC *Enterococcus* sp., *Bacteroides fragilis*, *Pseudomonas aeruginosa*,
CC *Klebsiella pneumoniae*, *Serratia marcescens*, *Mycobacterium tuberculosis*,
CC *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Haemophilus*
XX *Influenzae*, and *Staphylococcus saprophyticus*.

XX SQ Sequence 25 AA:

Alignment Scores:

Prod. No.:	744	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	21	Gaps:	0

US-09-835-992a-22 (1-843) x AAY69469 (1-25)

QY 96 TTAAAAACCTNGGMAA 113

Db 6 LeuLysAsnLeuGlyLys 11

Search completed: January 14, 2003, 17:27:28
Job time : 45.5252 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:21:54 ; Search time 11.5854 Seconds
(without alignments)
4281.862 Million cell updates/sec

Title: US-09-835-992A-22

Perfect score: 180
Sequence: 1 ggcacaaaattattctta.....taattccgaacataaagaag 843

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL-frame+ n2p.model -DEV=xlp
-O/cgn2.1/USPTO.spool/US09835992/runat.14012003.161527.4050/app.query.fasta.1.3932
-DB-Isued Patents.AA -QFMT-fastan -SUFFIX-olig.ra1 -MINMATCH=0.1 -LOORCL=0
-LOOEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-oligo -TRANS-human40.col
-LIST=45 -DOCLIN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE-LOCAL
-USFR=US09835992.qcgn.1.1.49.qrunat.14012003.161527.4050 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARN_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-YCAPOP=60 -YCAPEXT=60 -DELOP=6 -DELEXT=7

Database:

Issued_Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2.6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2.6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2.6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	3.3	35	2	US-08-142-551B-38
2	6	3.3	147	4	US-09-325-932A-113
3	6	3.3	164	3	US-08-765-381-2
4	6	3.3	176	2	US-08-619-542B-44
5	6	3.3	180	1	US-08-447-010-8
6	6	3.3	187	4	US-08-981-392-46
7	6	3.3	219	4	US-09-527-345-2
8	6	3.3	247	4	US-09-414-828A-14
9	6	3.3	257	2	US-08-451-947-10
10	6	3.3	257	2	US-08-424-826A-10
11	6	3.3	257	3	US-08-928-694-10
12	6	3.3	257	5	PCT-US91-05950-10

13	6	3.3	266	4	US-09-161-241-10	Sequence 10, Appl
14	6	3.3	273	2	US-08-701-935-5	Sequence 5, Appl1
15	6	3.3	273	3	US-09-134-591-5	Sequence 5, Appl1
16	6	3.3	291	4	US-09-420-786A-1	Sequence 1, Appl1
17	6	3.3	295	3	US-08-461-607-13	Sequence 13, Appl
18	6	3.3	295	4	US-09-363-600-13	Sequence 13, Appl
19	6	3.3	313	2	US-08-990-379-7	Sequence 7, Appl1
20	6	3.3	314	4	US-09-164-193-22	Sequence 22, Appl
21	6	3.3	319	4	US-09-221-448A-22	Sequence 22, Appl
22	6	3.3	319	4	US-09-100-664A-10	Sequence 10, Appl
23	6	3.3	319	4	US-09-335-983-10	Sequence 10, Appl
24	6	3.3	324	3	US-08-461-607-7	Sequence 7, Appl1
25	6	3.3	324	4	US-09-363-600-7	Sequence 7, Appl1
26	6	3.3	357	4	US-09-134-001C-3219	Sequence 3219, Ap
27	6	3.3	371	2	US-08-873-093-3	Sequence 2, Appl
28	6	3.3	390	2	US-08-873-093-3	Sequence 2, Appl
29	6	3.3	390	4	US-09-206-646-3	Sequence 3, Appl1
30	6	3.3	398	4	US-08-856-841-21	Sequence 21, Appl
31	6	3.3	443	1	US-08-342-411A-4	Sequence 4, Appl1
32	6	3.3	446	2	US-08-372-652-3	Sequence 3, Appl1
33	6	3.3	446	4	US-08-776-844-2	Sequence 2, Appl1
34	6	3.3	446	5	PCT-US95-16311-3	Sequence 3, Appl1
35	6	3.3	460	1	US-08-342-411A-2	Sequence 2, Appl1
36	6	3.3	461	1	US-08-350-518-2	Sequence 2, Appl1
37	6	3.3	461	1	US-08-350-283-2	Sequence 2, Appl1
38	6	3.3	461	2	US-08-646-248-2	Sequence 2, Appl1
39	6	3.3	461	5	PCT-US95-13924-2	Sequence 2, Appl1
40	6	3.3	461	5	PCT-US95-13931-2	Sequence 2, Appl1
41	6	3.3	464	1	US-08-969-630-5	Sequence 5, Appl1
42	6	3.3	470	1	US-08-514-014-2	Sequence 2, Appl1
43	6	3.3	470	2	US-08-833-823-3	Sequence 2, Appl1
44	6	3.3	486	4	US-09-134-001C-3593	Sequence 3593, Ap
45	6	3.3	504	4	US-09-126-420A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-142-551B-38
Sequence 38, Application US/08142551B
Patent No. 5814603
GENERAL INFORMATION:
APPLICANT: Oldenburg, Kevin R.
TITLE OF INVENTION: COMPOUNDS WITH PTH ACTIVITY AND
TITLE OF INVENTION: RECOMBINANT DNA VECTORS ENCODING SAME
NUMBER OF SEQUENCES: 132
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: 699 Prince Street
CITY: Alexandria
STATE: Virginia
COUNTRY: US
ZIP: 22313
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,551B
FILING DATE: 25-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,296
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/898,219
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/965,677
FILING DATE: 22-OCT-1992
ATTORNEY/AGENT INFORMATION:

NAME: Swiss, Gerald F.
REGISTRATION NUMBER: 30,113
REFERENCE/DOCKET NUMBER: 000324-010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 854-7400
TELEFAX: (415) 854-8275
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 35
OTHER INFORMATION: /note= "where 'Xaa' is selected
OTHER INFORMATION: from the group consisting of HoI, Ho, a homoserine
OTHER INFORMATION: amide, or the sequence of amino acids comprising
OTHER INFORMATION: residues 35-84 of PTH."
US-08-142-551B-38

Alignment Scores:
Pred. No.: 227 Length: 35
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-142-551B-38 (1-35)

QY 96 TTTAAACCTGAGAAA 113
|||||
DB 8 LeuLysAsnLeuGlyLys 13

RESULT 2

US-09-325-932A-173
Sequence 173, Application US/09325932A
Patent No. 6451604
GENERAL INFORMATION:
APPLICANT: Flinn, Barry
TITLE OF INVENTION: Compositions affecting programmed cell
TITLE OF INVENTION: death and their use in the modification of forestry plant develop
FILE REFERENCE: 1022
CURRENT APPLICATION NUMBER: US/09/325,932A
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 206
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 173
LENGTH: 147
TYPE: PRT
ORGANISM: Pinus radiata
US-09-325-932A-173

Alignment Scores:
Pred. No.: 180 Length: 147
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 4 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-325-932A-173 (1-147)

QY 627 CCCGTTAATCCCAAGA 610
|||||
DB 65 ProValLysSerGlnGly 70

RESULT 3
US-08-765-381-2
Sequence 2, Application US/08765381
Patent No. 6083724

GENERAL INFORMATION:
APPLICANT: Commonwealth Scientific and Industrial Research Organisation
TITLE OF INVENTION: No. 6083724el avian cytokines and genetic
TITLE OF INVENTION: sequences encoding same
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully Scott Murphy and Presser
STREET: 400 Garden City Plaza
CITY: Garden City, New York
STATE: New York
COUNTRY: UNITED STATES OF AMERICA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,381
FILING DATE: 19-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN1542/95
FILING DATE: 06-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00114
FILING DATE: 05-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: PRESSER, LEOPOLD
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-516-742-4343
TELEFAX: 1-516-742-4366
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-381-2

Alignment Scores:
Pred. No.: 176 Length: 164
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 3 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-765-381-2 (1-164)

QY 777 TTTAAACACACCTCC 794
|||||
DB 103 LeuLysAsnAsnLeuPro 108

RESULT 4

US-08-619-542B-44
Sequence 44, Application US/08619542B
Patent No. 5830662
GENERAL INFORMATION:
APPLICANT: The Trustees of Columbia University in the City
TITLE OF INVENTION: METHOD FOR CONSTRUCTION OF NORMALIZED
TITLE OF INVENTION: CDNA LIBRARIES
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/619,542B
 FILING DATE: June 21, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 42840-A-PCT-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 TELEX:
 INFORMATION FOR SEQ ID NO: 44:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 176 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-619-542B-44

Alignment Scores:
 Pred. No.: 174 Length: 176
 Score: 6.00 Matches: 6
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.33% Indels: 0
 DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-619-542B-44 (1-176)
 QY 767 GCTCCCGCCCTTAAAAA 784
 Db 110 GlyProProLeuLysLys 115

RESULT 5
 US-08-447-010-8
 Sequence 8, Application US/08447010
 Patent No. 5770718
 GENERAL INFORMATION:
 APPLICANT: MOEFATT, BARBARA
 TITLE OF INVENTION: GENE FOR APRT FROM PLANT TISSUE
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 330 University Avenue, Suite 701
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/447,010
 FILING DATE: 23-MAY-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/230,695
 FILING DATE: 21-APR-1994
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/888,132
 FILING DATE: 26-MAY-1992
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: STEWART, MICHAEL I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1811-183 MIS:vg
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 TELEX: 065-24567 SIMBAS
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 180 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-447-010-8

Alignment Scores:
 Pred. No.: 174 Length: 180
 Score: 6.00 Matches: 6
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.28% Indels: 0
 DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x US-08-447-010-8 (1-180)
 QY 771 GGACGNATCCCGTTTTC 754
 Db 169 GlyProIleProPhePhe 174

RESULT 6
 US-08-981-392-46
 Sequence 46, Application US/08981392
 Patent No. 6262025
 GENERAL INFORMATION:
 APPLICANT: Ish-Horowicz, David
 APPLICANT: Henrique, Domingos Manuel Pinto
 APPLICANT: Lewis, Julian Hart
 APPLICANT: Aravanis-Tsakonas, Spyridon
 TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES
 TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON
 NUMBER OF SEQUENCES: 94
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennle & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10036/2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/981,392
 FILING DATE: 22-DEC-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Antler, Adriane M.
 REGISTRATION NUMBER: 32,605
 REFERENCE/DOCKET NUMBER: 7326-038
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 187 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-981-392-46

Alignment Scores:

Pred. No.: 173 Length: 187
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-22 (1-843) x US-08-981-392-46 (1-187)

OY 777 AGGGGGGACGACNATCCCG 760

DB 3 ArgGlyGlyProIlePro 8

RESULT 7

US-09-527-345-2
Sequence 2, Application US/09527345
Patent No. 6331413

GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.

APPLICANT: Adler, David A.

TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71

CURRENT APPLICATION NUMBER: US/09/527,345
CURRENT FILING DATE: 1999-03-17

PRIOR APPLICATION NUMBER: US 60/124,820
PRIOR FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 219

TYPE: PRT

ORGANISM: Homo sapiens

US-09-527-345-2

Alignment Scores:
Pred. No.: 168 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-527-345-2 (1-219)

OY 786 AACCTCCCCCNCCTTGTG 803

DB 46 AsnLeuProProIleu 51

RESULT 8

US-09-414-828a-14
Sequence 14, Application US/09414828a
Patent No. 6300084

GENERAL INFORMATION:
APPLICANT: Drubbin, David G.

APPLICANT: Hofmann, Christian J.

TITLE OF INVENTION: Anti-Mitotic Agents and Processes
FILE REFERENCE: Mitotic Spindle Proteins

CURRENT APPLICATION NUMBER: US/09/414,828a
CURRENT FILING DATE: 1999-10-06

PRIOR APPLICATION NUMBER: 60/103,684
PRIOR FILING DATE: 1998-10-08

NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent Ver. 2.0

SEQ ID NO 14

LENGTH: 247

TYPE: PRT

ORGANISM: Saccharomyces cerevisiae

US-09-414-828a-14

Alignment Scores:
Pred. No.: 165 Length: 247
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-414-828a-14 (1-247)

OY 777 TTTAAACACACCTTCCC 794

DB 85 LeuLysAsnAsnLeuPro 90

RESULT 9

US-08-451-947-10
Sequence 10, Application US/08451947
Patent No. 5702906

GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.

APPLICANT: ROSENTHAL, ARNON

TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd
CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/451,947
FILING DATE:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/426419
FILING DATE: 19-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/030013

FILING DATE: 22-MAR-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/648482
FILING DATE: 31-JAN

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707

FILING DATE: 1991
ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: 666P2C1D2
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881

TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids

TYPE: amino acid
TOPOLOGY: linear

US-08-451-947-10

Alignment Scores:
Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x US-08-451-947-10 (1-257)

OY 378 ATTTTTCCTCCCAATCA 361

|||||
DB 26 llephepseroanspro 31
RESULT 10
US-08-424-826A-10
Sequence 10, Application US/08424826A
Patent No. 5830858
GENERAL INFORMATION:
APPLICANT: Rosenthal, Arnon
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/424,826A
FILING DATE: 19-Apr-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/240387
FILING DATE: 10-May-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
FILING DATE: 31-Jan-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/587707
FILING DATE: 25-Sep-1990
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Phd., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0666P1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-424-826A-10
Alignment Scores:
Pred. No.: 164
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.28%
DB: 2
Gaps: 0
US-09-835-992a-22 (1-843) x US-08-424-826A-10 (1-257)
Qy 378 ATTTTTCCTCCCAATCA 361
|||||
DB 26 llephepseroanspro 31
RESULT 11
US-08-928-694-10
Sequence 10, Application US/08928694
Patent No. 6037320
GENERAL INFORMATION:
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 1 DNA Way
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WinPatIn (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/928,694
;; FILING DATE: 12-Sep-1997
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/451947
;; FILING DATE: 26-May-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/426419
;; FILING DATE: 19-Apr-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/030013
;; FILING DATE: 22-Mar-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/648482
;; FILING DATE: 31-Jan-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/587707
;; FILING DATE: 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Torchia, Phd., Timothy E.
;; REGISTRATION NUMBER: 36,700
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650/225-8674
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 257 amino acids
;; TYPE: Amino Acid
;; TOPOLOGY: Linear
;; US-08-928-694-10
Alignment Scores:
Pred. No.: 164
Score: 6.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.28%
DB: 3
Gaps: 0
US-09-835-992a-22 (1-843) x US-08-928-694-10 (1-257)
Qy 378 ATTTTTCCTCCCAATCA 361
|||||
DB 26 llephepseroanspro 31
RESULT 12
PCT-US91-06950-10
Sequence 10, Application PC/TUS9106950
GENERAL INFORMATION:
APPLICANT: GENENTECH, INC.
APPLICANT: ROSENTHAL, ARNON
TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 Inch, 360 Kb Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/06950
FILING DATE: 19910924
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/648482
PRIOR APPLICATION DATA: 07/587707
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 666P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-06950-10

Alignment Scores:
Pred. No.: 164 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 5 Gaps: 0

US-09-835-992a-22 (1-843) x PCT-US91-06950-10 (1-257)

OY 378 ATTTTTCCTCCCATCA 361
|||||
Db 26 IllePheProAsnPro 31

RESULT 13
US-09-161-241-10
Sequence 10, Application US/09161241
Patent No. 6344541
GENERAL INFORMATION:
APPLICANT: Bass, Michael B
APPLICANT: Sullivan, John K
APPLICANT: Theill, Lars E
APPLICANT: Wang, Daquan
TITLE OF INVENTION: NOVEL DKR POLYPEPTIDES
FILE REFERENCE: A-548
CURRENT APPLICATION NUMBER: US/09/161,241
CURRENT FILING DATE: 1998-09-25
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 266
TYPE: PRT
ORGANISM: Human
US-09-161-241-10

Alignment Scores:
Pred. No.: 163 Length: 266
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 4 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-161-241-10 (1-266)

OY 786 AACCTTCCCCCNCCTTG 803
|||||
Db 44 AsnLeuProProLeu 49

RESULT 14
US-08-701-935-5
Sequence 5, Application US/08701935
Patent No. 5843683
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Au-Yang, Janice
APPLICANT: Hillman, Jennifer
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PROTEASOME
SUBUNIT PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA: 1.5
APPLICATION NUMBER: US/08/701,935
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0119 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 307307
US-08-701-935-5

Alignment Scores:
Pred. No.: 162 Length: 273
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x US-08-701-935-5 (1-273)

OY 617 GATTTACGGGTCCCAA 634
|||||
Db 143 AspleuThrGlyProGln 148

RESULT 15
US-09-134-591-5
Sequence 5, Application US/09134591
Patent No. 6018028
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Au-Yang, Janice
APPLICANT: Hillman, Jennifer

APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN PROTEASOME
SUBUNIT PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,591
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/701,935
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0119 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 307307
US-09-134-591-5

Alignment Scores:
Pred. No.: 162 Length: 273
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 3 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-134-591-5 (1-273)

OY 617 GATTACGGGTCCCA 634
|||||
Db 143 AspleurthrcglyProgin 148

Search completed: January 14, 2003, 17:42:19
Job time : 14.8354 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleole - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:29:25 ; Search time 8.07072 Seconds
(without alignments)
4052.943 Million cell updates/sec

Title: US-09-835-992A-22
Perfect score: 180
Sequence: 1 ggcacaaataattattna.....taaatcgacaataaaag 843

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 118974 seqs, 19401057 residues

Word size: 1

Total number of hits satisfying chosen parameters: 214280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p.model -DEV=xlp
-G/cgn2_1/USPNC_spool/US09835983/runat_14012003_161528_4118/app_query.fasta.1.3932
-DB=Published_Applications_MA -QFMT=fastan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRAN=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09835992.@CGN.1.1.28 @runat_14012003_161528_4118
-NCPu=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

- 1: /cgnt2_6/ptodata/2/pubpa/US08_NEW_PUB.pep.*
- 2: /cgnt2_6/ptodata/2/pubpa/pct_NEW_PUB.pep.*
- 3: /cgnt2_6/ptodata/2/pubpa/US06_NEW_PUB.pep.*
- 4: /cgnt2_6/ptodata/2/pubpa/US06_PUBCOMB.pep.*
- 5: /cgnt2_6/ptodata/2/pubpa/US07_NEW_PUB.pep.*
- 6: /cgnt2_6/ptodata/2/pubpa/US07_PUBCOMB.pep.*
- 7: /cgnt2_6/ptodata/2/pubpa/pctUS_PUBCOMB.pep.*
- 8: /cgnt2_6/ptodata/2/pubpa/US08_PUBCOMB.pep.*
- 9: /cgnt2_6/ptodata/2/pubpa/US09_NEW_PUB.pep.*
- 10: /cgnt2_6/ptodata/2/pubpa/US09_PUBCOMB.pep.*
- 11: /cgnt2_6/ptodata/2/pubpa/US10_NEW_PUB.pep.*
- 12: /cgnt2_6/ptodata/2/pubpa/US10_PUBCOMB.pep.*
- 13: /cgnt2_6/ptodata/2/pubpa/US60_NEW_PUB.pep.*
- 14: /cgnt2_6/ptodata/2/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
C 1	6	3.3	46	US-09-864-761-47474
C 2	6	3.3	74	US-09-764-847-677
C 3	6	3.3	91	US-09-764-877-1669
C 4	6	3.3	129	US-09-867-550-1822

5	6	3.3	140	10	US-09-893-737-118	Sequence 118, App
6	6	3.3	140	10	US-09-746-491-6	Sequence 6, Appl
7	6	3.3	182	10	US-09-764-903-47	Sequence 47, Appl
C 8	6	3.3	187	10	US-09-908-323-46	Sequence 46, Appl
9	6	3.3	191	9	US-09-738-626-5878	Sequence 5878, Ap
10	6	3.3	219	9	US-09-922-469-2	Sequence 2, Appl1
11	6	3.3	219	10	US-09-922-480-2	Sequence 2, Appl1
C 13	6	3.3	219	10	US-09-923-236-2	Sequence 10, Appl
C 14	6	3.3	257	8	US-08-450-842-10	Sequence 10, Appl
15	6	3.3	259	9	US-09-996-634-133	Sequence 133, App
16	6	3.3	266	9	US-09-974-298-89	Sequence 89, Appl
17	6	3.3	266	9	US-09-978-2954-456	Sequence 456, App
18	6	3.3	266	9	US-09-976-736-10	Sequence 428, App
19	6	3.3	266	9	US-09-978-697-456	Sequence 456, App
20	6	3.3	266	9	US-09-978-1924-456	Sequence 456, App
21	6	3.3	266	9	US-09-999-8324-456	Sequence 456, App
22	6	3.3	266	9	US-09-978-189-456	Sequence 456, App
C 23	6	3.3	266	9	US-10-028-072-428	Sequence 428, App
C 24	6	3.3	297	10	US-09-925-297-635	Sequence 635, App
C 25	6	3.3	351	9	US-09-895-9134-170	Sequence 170, App
26	6	3.3	373	9	US-09-738-626-6663	Sequence 6663, Ap
27	6	3.3	390	9	US-10-137-731-2	Sequence 2, Appl1
28	6	3.3	409	10	US-10-072-130-3	Sequence 3, Appl1
29	6	3.3	437	10	US-09-790-264-42	Sequence 42, Appl
C 30	6	3.3	437	10	US-09-790-264-30	Sequence 30, Appl
C 31	6	3.3	446	10	US-09-909-446-2	Sequence 2, Appl1
C 32	6	3.3	446	10	US-09-909-325-2	Sequence 2, Appl1
C 33	6	3.3	446	10	US-09-909-325-2	Sequence 2, Appl1
C 34	6	3.3	446	12	US-10-013-823-2	Sequence 3686, Ap
C 35	6	3.3	461	12	US-10-013-823-3	Sequence 3, Appl1
36	6	3.3	470	9	US-10-114-893-8	Sequence 8, Appl1
37	6	3.3	480	10	US-09-790-264-32	Sequence 32, Appl
38	6	3.3	482	10	US-09-796-856-40	Sequence 40, Appl
39	6	3.3	533	10	US-09-836-410A-1	Sequence 1, Appl1
C 40	6	3.3	533	10	US-09-836-503B-2	Sequence 2, Appl1
C 41	6	3.3	630	10	US-09-843-598-10	Sequence 5, Appl
42	6	3.3	700	10	US-09-836-503B-5	Sequence 4, Appl1
43	6	3.3	710	10	US-09-836-503B-4	Sequence 3, Appl1
44	6	3.3	723	10	US-09-836-503B-3	Sequence 3, Appl1
C 45	6	3.3	748	10	US-09-864-761-43244	Sequence 43244, A

ALIGNMENTS

RESULT 1
US-09-864-761-47474
Sequence 47474, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenshang
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmtca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 47474
;; LENGTH: 46
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC022325.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;; OTHER INFORMATION: EST_HUMAN HIT: AV681548.1, EVALU8 5.20e+00
;; OTHER INFORMATION: SWISSPROT HIT: O76039, EVALU8 8.20e+00
US-09-864-761-47474

Alignment Scores:
Pred. No.: 161          Length: 46
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 10                  Gaps: 0

US-09-835-992a-22 (1-843) x US-09-864-761-47474 (1-46)
QY 802 AAGGNGGGGGAAGCTTG 785
   |||||||
Db 39 LysGlyGlyGlyArgLeu 44

RESULT 2
US-09-764-847-677
;; Sequence 677, Application US/09764847
;; Patent No. US20020132767A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC009
;; CURRENT APPLICATION NUMBER: US/09/764,847
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - consult PAM or file wrapper
;; NUMBER OF SEQ ID NOS: 2003
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 677
;; LENGTH: 74
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (33)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (53)
```

```
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (55)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-677

Alignment Scores:
Pred. No.: 145          Length: 74
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 10                  Gaps: 0

US-09-835-992a-22 (1-843) x US-09-764-847-677 (1-74)
QY 376 TTTTTCCTCCCATCCAA 359
   |||||||
Db 61 PhepHeProGlnSerIys 66

RESULT 3
US-09-764-877-1669
;; Sequence 1669, Application US/09764877
;; Patent No. US20020147140A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PC005
;; CURRENT APPLICATION NUMBER: US/09/764,877
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - refer to PAM or file wrapper
;; NUMBER OF SEQ ID NOS: 4031
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1669
;; LENGTH: 91
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (12)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (15)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1669

Alignment Scores:
Pred. No.: 139          Length: 91
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 10                  Gaps: 0

US-09-835-992a-22 (1-843) x US-09-764-877-1669 (1-91)
QY 777 AGGGGGGAGCNCATCCG 760
   |||||||
Db 61 ArgGlyGlyProIlePro 66

RESULT 4
US-09-867-550-1822
;; Sequence 1822, Application US/09867550
;; Patent No. US20020082206A1
;; GENERAL INFORMATION:
;; APPLICANT: Leach, Martin D.
;; APPLICANT: Menhaban, Foad,
;; APPLICANT: Conley, Pamela
;; APPLICANT: Law, Debbie
;; APPLICANT: Topper, James
;; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
;; TITLE OF INVENTION: Thereby
;; FILE REFERENCE: 21402-013 (Cura-313)
```



```
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1822
; LENGTH: 129
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (72)
; OTHER INFORMATION: wherein Xaa may be any one of Arg or Cys or Gly or Ser
; NAME/KEY: VARIANT
; LOCATION: (92)
; OTHER INFORMATION: wherein Xaa may be any one of Gln or His
US-09-867-550-1822

Alignment Scores:
Pred. No.: 129          Length: 129
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 10                 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-867-550-1822 (1-129)

QY 635 ATTGGGACCCCTTAAAT 618
Db 121 llllllllllllllllll 126

RESULT 5
US-09-893-737-118
; Sequence 118, Application US/09893737
; Patent No. US2002010853A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 118
; LENGTH: 140
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-893-737-118

Alignment Scores:
Pred. No.: 127          Length: 140
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33%      Indels: 0
DB: 10                 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-893-737-118 (1-140)

QY 789 CTTCCCCNCCTTTGGCC 806
Db 11 llllllllllllllllll 16

RESULT 6
US-09-746-491-6
; Sequence 6, Application US/09746491
; Patent No. US20020137202A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: NO. US20020137202A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-746-491-6

Alignment Scores:
Pred. No.: 127          Length: 140
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33%      Indels: 0
DB: 10                 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-746-491-6 (1-140)

QY 789 CTTCCCCNCCTTTGGCC 806
Db 11 llllllllllllllllll 16

RESULT 7
US-09-764-903-47
; Sequence 47, Application US/09764903
; Patent No. US20020090674A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antbodies
; FILE REFERENCE: Ptz28
; CURRENT APPLICATION NUMBER: US/09/764,903
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47
; LENGTH: 182
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-764-903-47

Alignment Scores:
Pred. No.: 120          Length: 182
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33%      Indels: 0
DB: 10                 Gaps: 0

US-09-835-992A-22 (1-843) x US-09-764-903-47 (1-182)

QY 792 CCCCCNCCTTTGGCCAG 809
Db 123 ProProProLeuAlaGln 128

RESULT 8
US-09-908-322-46
; Sequence 46, Application US/09908322
; Patent No. US20020107194A1
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; TITLE OF INVENTION:
; Lewis, Julian Hart
; Artavanis-Tsakonas, Spyridon
; Gray, Grace
```

VERTEBRATE DELTA GENE AND METHODS BASED THEREON

NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036/2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/908,322
FILING DATE: 17-Jul-2001
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/981,392
FILING DATE: 22-DEC-1997

ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: amino acid
STRANDEDNESS: <unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-908-322-46

Alignment Scores:
Pred. No.: 119 Length: 187
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-908-322-46 (1-187)

QY 777 AGGGGGGACGACNATCCCG 760
Db 3 ArgglyglyProIlePro 8

RESULT 9
US-09-738-626-5878
Sequence 5878, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5878
LENGTH: 191
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5878

Alignment Scores:
Pred. No.: 119 Length: 191
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-738-626-5878 (1-191)

QY 616 GGATTTAAGGGGTCCCGCA 633
Db 72 GlyPheAsnglySerPro 77

RESULT 10
US-09-922-469-2
Sequence 2, Application US/09922469
Patent No. US20020173027A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul A.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,469
CURRENT FILING DATE: 2001-08-03
PRIOR FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-469-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 9 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-922-469-2 (1-219)

QY 786 AACCTTCCCGCCGCTTTG 803
Db 46 AsnLeuProProIleu 51

RESULT 11
US-09-922-480-2
Sequence 2, Application US/09922480
Patent No. US20020081701A1
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: SECRETED SALIVARY ZSIG63 POLYPEPTIDE
FILE REFERENCE: 97-71
CURRENT APPLICATION NUMBER: US/09/922,480
CURRENT FILING DATE: 2001-08-03

;; PRIOR APPLICATION NUMBER: US 60/124,820
;; PRIOR FILING DATE: 1999-03-17
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FASTSEQ for Windows Version 3.0
;; SEQ ID NO 2
;; LENGTH: 219
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-922-480-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-922-480-2 (1-219)

OY 786 AACCTTCCCCCNCCTTGG 803
Db 46 AsnLeuProProProlou 51

RESULT 12
US-09-923-236-2
; Sequence 2, Application US/09923236
; Patent No. US20020090677A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Adler, David A.
; TITLE OF INVENTION: SECRETED SALIVARY ZS1G63 POLYPEPTIDE
; FILE REFERENCE: 97-71
; CURRENT APPLICATION NUMBER: US/09/923,236
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/124,820
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 219
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-923-236-2

Alignment Scores:
Pred. No.: 115 Length: 219
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x US-09-923-236-2 (1-219)

OY 786 AACCTTCCCCCNCCTTGG 803
Db 46 AsnLeuProProProlou 51

RESULT 13
US-08-450-842-10
; Sequence 10, Application US/08450842
; Patent No. US20020045576A1
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; APPLICANT: ROSENTHAL, ARNON
; TITLE OF INVENTION: NOVEL NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: palin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/450,842
;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/426419
;; FILING DATE: 19-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/030013
;; FILING DATE: 22-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/648482
;; FILING DATE: 31-JAN
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/587707
;; FILING DATE: 1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Torchia, Timothy E.
;; REGISTRATION NUMBER: 36,700
;; REFERENCE/DOCKET NUMBER: 666P2CID3
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-8674
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 257 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
US-08-450-842-10

Alignment Scores:
Pred. No.: 111 Length: 257
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 8 Gaps: 0

US-09-835-992a-22 (1-843) x US-08-450-842-10 (1-257)

OY 378 ATTTTTCCTCCCAATCA 361
Db 26 IlePhePheProAsnPro 31

RESULT 14
US-09-996-634-133
; Sequence 133, Application US/09996634
; Patent No. US20020172684A1
; GENERAL INFORMATION:
; APPLICANT: Nano, Francis
; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
; FILE REFERENCE: 61260
; CURRENT APPLICATION NUMBER: US/09/996,634
; CURRENT FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: 09/447,135
; PRIOR FILING DATE: 2000-01-03
; PRIOR APPLICATION NUMBER: 08/990,823
; PRIOR FILING DATE: 1997-12-15
; PRIOR APPLICATION NUMBER: US 96/10375
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: 60/000,254
; PRIOR FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: Patent Ver. 2.0

```

; SEQ ID NO 133
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-996-634-133

```

```

Alignment Scores:
Pred. No.: 111          Length: 259
Score: 6.00           Matches: 6
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 3.28%           Indels: 0
DB: 9                      Gaps: 0

```

```

US-09-835-992A-22 (1-843) x US-09-996-634-133 (1-259)

```

```

QY 802 AAAGNGGGGAAGTTG 785
DB 204 LysGlyGlyArgLeu 209

```

RESULT 15

```

US-09-974-298-89
; Sequence 89, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 89
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 2401150CD1
US-09-974-298-89

```

```

Alignment Scores:
Pred. No.: 111          Length: 266
Score: 6.00           Matches: 6
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 3.33%           Indels: 0
DB: 9                      Gaps: 0

```

```

US-09-835-992A-22 (1-843) x US-09-974-298-89 (1-266)

```

```

QY 786 AACCTTCCCCGNCCTTG 803
DB 44 AsnLeuProProLeu 49

```

```

Search completed: January 14, 2003, 18:08:22
Job time : 10.0707 secs

```


C:Species: *Triticum aestivum* (common wheat)
C:Date: 13-Jan-1995 #sequence_rev13-Jan-1995 #text_change 29-Oct-1999
C:Accession: S44154
R:Kroczyńska, B.; Buchowicz, J.
submitted to the EMBL Data Library, November 1993
A:Reference number: S44153
A:Accession: S44154
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1786 <KRO>
A:Cross-references: EMBL:X76181; NID:9473172; PIDN:CAA53774.1; PID:9473173

Alignment Scores:
Pred. No.: 27.9 Length: 86
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.83% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x S44154 (1-86)

QY 799 GGNGGGGAGAGCTTGTCTT 779
Db 37 GlyGlyGlyArgLeuPhePhe 43
|||||
|

RESULT 3
A:70864
hypothetical protein RV2450c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_rev17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70864
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70864
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-172 <COL>
A:Cross-references: GB:AL021246; GB:AL123456; NID:93261507; PIDN:CAA16027.1; PID:el373755
C:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2450c

Alignment Scores:
Pred. No.: 24.7 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
Gaps: 0

US-09-835-992A-22 (1-843) x A70864 (1-172)

QY 786 AACCTTCCCCNCCCTTGGCC 806
Db 73 AsnLeuProProLeuAla 79
|||||
|

RESULT 4
A:690126
hypothetical protein orf223 [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_rev10-May-2001 #text_change 24-May-2001
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reiter,
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671

[illegible]

A:Molecule type: DNA
A:Residues: 1-387 <HBN>
A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192711; PID:AA019064.1
C:Genetics:
A:Note: gene 28
C:Superfamily: sopa protein

Alignment Scores:
Pred. No.: 21.3 Length: 387
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x T13114 (1-387)

OY 670 TTACATTTTACCTTGTAAAG 690
DB 169 LeuProPheTyrLeuGlyLys 175

RESULT 7
68865
fimbrial protein h1fe precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
C>Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 08-Oct-1999
C:Accession: S68865; S54431
R:van Ham, M.S.
Submitted to the EMBL Data Library, May 1994
A:Reference number: S68865
A:Accession: S68865
A:Molecule type: DNA
A:Residues: 1-435 <VAN>
A:Cross-references: EMBL:Z33502; NID:9535165; PID:CAA83904.1; PID:9535169
R:van Ham, S.M.; van Alphen, L.; Mool, F.R.; van Putten, J.P.M.
Mol. Microbiol. 13, 673-684, 1994
A:Title: The fimbrial gene cluster of Haemophilus influenzae type b.
A:Reference number: S54428; MUID:95089703; PMID:7997179
A:Accession: S54431
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 406-435 <VAN>
A:Cross-references: EMBL:Z33502
C:Genetics:
A:Gene: h1fe

Alignment Scores:
Pred. No.: 20.9 Length: 435
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x S68865 (1-435)

OY 776 CTTAAAAACAACTTCCCC 796
DB 389 LeuYelYsgInProSerPro 395

RESULT 8
747919
hypothetical protein T20K12.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Jun-2000
C:Accession: T47919
R:De Haan, M.; Maarse, A.C.; Grivel, L.A.; Mewes, H.W.; Lemcke, K.; Meyer, K.F.X.; Quer
Submitted to the Protein Sequence Database, January 2000
A:Reference number: Z24480
A:Accession: T47919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-498 <DEH>

A:Cross-references: EMBL:AL137898
A:Experimental source: cultivar Columbia; BAC clone T20K12
C:Genetics:
A:Map position: 3
A:Introns: 46/3; 222/2; 286/3; 308/2; 378/3
A:Note: T20K12.170
C:Superfamily: Arabidopsis thaliana hypothetical protein T20K12.190

Alignment Scores:
Pred. No.: 20.4 Length: 498
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x T47919 (1-498)

OY 817 TCCGCTTAAATCCGACAAAT 837
DB 84 SerArgLeuAsnProAsnAsn 90

RESULT 9
748800
SMT4 related protein [Imported] - Neurospora crassa
N:Alternate names: protein 15E6.80
C:Species: Neurospora crassa
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: T48800
R:Schulte, U.; Algen, V.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu
Submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24541
A:Accession: T48800
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1240 <SCH>
A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.80
A:Experimental source: cosmid contig 15E6; strain 74
C:Genetics:
A:Gene: NCSP:15E6.80
A:Map position: 2
A:Introns: 8/3; 358/2

Alignment Scores:
Pred. No.: 17.3 Length: 1240
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x T48800 (1-1240)

OY 767 GGTCCCCCTTAAAAACAA 787
DB 29 GlyProLeuYsgInLysGln 35

RESULT 10
VCBEN7
major capsid protein - saimiriine herpesvirus 1 (strain 11)
C:Species: saimiriine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
R:Albrecht, J.
Submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36806
A:Accession: E36808
A:Molecule type: DNA
A:Residues: 1-1371 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PID:CAA45648.1; PID:g60346
R:Albrecht, J.C.; Nichols, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.;

J. Virol. 66, 5047-5058, 1992
A:Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688; PMID:132187
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Genetics:
A:Gene: 25
C:Superfamily: varicella-zoster virus major capsid protein
C:Keywords: capsid protein

Alignment Scores:
Pred. No.: 17 Length: 1371
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992A-22 (1-843) x VCBEM7 (1-1371)

OY 786 AACCTTCCCCCNCCTTTGGCC 806
|||||
DB 560 AsnleuProProProleuAla 566

RESULT 11
T42938
major capsid protein - ateline herpesvirus 3 (strain 73)
C:Species: ateline herpesvirus 3
A:Variety: strain 73
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T42938
R:Albrecht, J.C.; Fleckenstein, B.
Submitted to the EMBL Data Library, August 1998
A:Description: Primary structure of the herpesvirus ateles genome.
A:Reference number: 222274
A:Accession: T42938
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1371 <ALB>
A:Cross-References: EMBL:AF083424; PIDN:AAC95549.1
A:Experimental source: strain 73
C:Superfamily: varicella-zoster virus major capsid protein

Alignment Scores:
Pred. No.: 17 Length: 1371
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x T42938 (1-1371)

OY 786 AACCTTCCCCCNCCTTTGGCC 806
|||||
DB 560 AsnleuProProProleuAla 566

RESULT 12
T07328
hypothetical protein 60c - Chlorella vulgaris chloroplast
C:Species: chloroplast Chlorella vulgaris
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C:Accession: T07328
R:Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakas
Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997
A:Title: Complete nucleotide sequence of the chloroplast genome from the green alga Chlo
A:Reference number: 215985; MUID:97303241; PMID:9159184
A:Accession: T07328
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-60 <WAK>
A:Cross-References: EMBL:AB001684; NID:g2224352; PIDN:BA457976.1; PID:g2224492
C:Genetics:

A:Genome: chloroplast
C:Keywords: chloroplast

Alignment Scores:
Pred. No.: 426 Length: 60
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x T07328 (1-60)

OY 784 TTTTAAAGGGGGGAGACC 767
|||||
DB 28 PhepheylsGlyGlythr 33

RESULT 13
F81945
probable hemeolysin NMA0983 [imported] - Neisseria meningitidis (strain 22491 serogro
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: F81945
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: F81945
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-68 <PAR>
A:Cross-References: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84253.1; PID:g737
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMA0983

Alignment Scores:
Pred. No.: 417 Length: 68
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x F81945 (1-68)

OY 805 GCCAAGGNGGGGAGAG 788
|||||
DB 31 AlaYsGlyGlyGlyVar 36

RESULT 14
C83058
hypothetical protein PA4703 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83058
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83058
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-90 <STO>
A:Cross-References: GB:AE004684; GB:AE004091; NID:g9950954; PIDN:AAG08089.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4703

Alignment Scores:

Pred. No.: 396 Length: 90
 Score: 6.00 Matches: 6
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.33% Indels: 0
 DB: 2 Gaps: 0

US-09-835-992A-22 (1-843) x C83058 (1-90)

OY 96 TTAATAAACCTNGGAAA 113
 |||||
 Db 8 LeuLysAnLeuGLyLys 13

RESULT 15

T27091

hypothetical protein Y51B9A.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27091

R:Barlow, K.

submitted to the EMBL Data Library, June 1998

A:Reference number: 220308

A:Accession: T27091

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-90 <Wild>

A:Cross-references: EMBL:AL023845; PIDN:CA19540.1; GSPDB:GN00020; CESP:Y51B9A.5

A:Experimental source: clone Y51B9A

C:Genetics:

A:Gene: CESP:Y51B9A.5

A:Map position: 2

A:Introns: 21/3

Alignment Scores:

Pred. No.:	396	Length:	90
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.33%	Indels:	0
DB:	2	Gaps:	0

US-09-835-992A-22 (1-843) x T27091 (1-90)

OY 792 CCCCNCCTTGGCCAC 809
 |||||
 Db 45 ProProProlenAlaGln 50

Search completed: January 14, 2003, 17:40:51
 Job time : 34.2475 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 14, 2003, 17:15:54 ; Search time 13.0173 Seconds
(without alignments)
5372.014 Million cell updates/sec

Title: US-09-835-992a-22

Perfect score: 180
Sequence: 1 ggcacaaataattatlna.....taalcgcacataaagaag 843

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp
-O=/gen2.1/USPRO.spool/US0983599/runat.14012003.161525.4014/app_query.fasta.1.3932
-DB=SwissProt_40 -QPM=fastan -SUFFIX=oligo -rnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09835992.ecgn.1.1.77.ernat.14012003.161525.4014 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -MGC_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7
-Ygapop=60 -Ygapext=60 -DELop=6 -DELExt=7

Database : SwissProt_40.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match Length	ID	Description
1	7	3.9	337	1 RPOA_CAMJ
2	7	3.9	435	1 HFEI_HAEIN
3	7	3.9	1371	1 VCAP_HSVSA
4	6	3.3	92	1 Y031_TREPA
5	6	3.3	164	1 ING_CHICK
6	6	3.3	164	1 ING_CHICK
7	6	3.3	164	1 ING_MELGA
8	6	3.3	169	1 ING_PIRAC
9	6	3.3	180	1 RLG_CHLPN
10	6	3.3	180	1 APT_MASHI
11	6	3.3	180	1 APT_MOUSE
12	6	3.3	180	1 APT_MOUSE
13	6	3.3	180	1 APT_MOUSE
14	6	3.3	183	1 APT_MOUSE
15	6	3.3	218	1 RRF_DEIRA
16	6	3.3	226	1 YPDC_BACSU
17	6	3.3	227	1 YRAL_YEAST
18	6	3.3	238	1 TPIS_SULSO
19	6	3.3	239	1 TCDE_PSESO
20	6	3.3	239	1 PNDG_SALTY

C 19	6	3.3	241	1 PSM_SULSO	09ux6 sulfolobus
C 20	6	3.3	241	1 YG1_YEAS	P53168 saccharomyc
C 21	6	3.3	254	1 HIS6_LEPBO	092gm0 leptospira
C 22	6	3.3	254	1 HIS6_LEPIN	094h7 leptospira
C 23	6	3.3	257	1 NTS6_HUMAN	P34133 homo sapien
C 24	6	3.3	251	1 LPRF_MYCTU	P71788 mycobacteri
C 25	6	3.3	263	1 DNAA_SPLAP	P35892 splioplasm
C 26	6	3.3	266	1 DK1_HUMAN	094907 homo sapien
C 27	6	3.3	269	1 PR11_YEAS	007350 saccharomyc
C 28	6	3.3	269	1 HXA9_MORSA	09p45 morone saxa
C 29	6	3.3	269	1 THID_TREPA	083153 treponema p
C 30	6	3.3	272	1 DK1_MOUSE	054908 mus musculu
C 31	6	3.3	273	1 PSBA_HUMAN	P40306 homo sapien
C 32	6	3.3	284	1 RS2_MYCE	P47316 mycoplasma
C 33	6	3.3	290	1 YAS7_METJA	058475 methanococc
C 34	6	3.3	298	1 Y295_RICPR	092d21 rickettsia
C 35	6	3.3	298	1 YDHF_ECOLI	P76187 escherichia
C 36	6	3.3	301	1 MCAT_HUMAN	043772 homo sapien
C 37	6	3.3	301	1 MCAT_MOUSE	09226 mus musculu
C 38	6	3.3	307	1 K1PF_BORBU	051575 borrelia bu
C 39	6	3.3	316	1 Y015_CAEEL	009278 caenorhabd
C 40	6	3.3	318	1 DUS2_MOUSE	005922 mus musculu
C 41	6	3.3	320	1 Y149_MYCPN	P75583 mycoplasma
C 42	6	3.3	326	1 NDM1_CHOCR	P48858 chondrus cf
C 43	6	3.3	326	1 RPOA_FUSNN	08re44 fusobacteri
C 44	6	3.3	332	1 ETFA_CAEEL	093615 caenorhabd
C 45	6	3.3	332	1 G3P_BUCAL	P57384 buchnera ap

ALIGNMENTS

RESULT 1

ID RPOA_CAMJ STANDARD; PRT; 337 AA.

AC 09PM80;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase alpha chain (EC 2.7.7.6) (Transcriptase
DE alpha chain) (RNA polymerase alpha subunit).

GN RPOA OR CUI595.

OS Campylobacter jejuni.

OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;

CC Campylobacter.

OX NCBI_TaxID:197;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE-NCIC 11168;

RX MEDLINE-20150912; PubMed-10668204;

RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,

RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,

RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,

RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,

RA Whitehead S., Barrett B.G.;

RT "The genome sequence of the food-borne pathogen Campylobacter jejuni

RT reveals hypervariable sequences.";

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS

CC SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +

CC (RNA)(N).

CC -1- SUBUNIT: CONSISTS OF A SIGMA FACTOR AND THE RNAP CORE ENZYME WHICH

CC IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, 1 BETA' CHAIN AND 1

CC OMEGA CHAIN (BY SIMILARITY).

CC -1- DOMAIN: THE AMINO-TERMINAL PORTION IS INVOLVED IN THE ASSEMBLY OF

CC CORE RNAP, WHEREAS THE C-TERMINAL IS INVOLVED IN INTERACTION WITH

CC TRANSCRIPTIONAL REGULATORS (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE ALPHA CHAIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AL139079; CAB73583.1; -
DR HSSP: P00574; ICOO.
DR InterPro: IPR001700; RNA_pol_A_bac.
DR Pfam: PF01000; RNA_pol_A_bac; 1.
DR Pfam: PF03118; RNA_pol_A_CTD; 1.
DR ProDom: PD001179; RNA_pol_A_bac; 1.
KW Transferase: Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 337 AA; 37686 MW; 4BD2F59205DDE143 CRC64;

Alignment Scores:
Pred. No.: 10.8 Length: 337
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x RPOA_CAMJE (1-337)
QY 96 TTTAAAAACCTNGAAAAAAA 116
DB 290 LeuLysAsnLeuLysLysLys 296

RESULT 2
HFE1_HAEIN
ID HFE1_HAEIN STANDARD; PRT; 435 AA.
AC P45994;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Minor fibrial subunit hife precursor.
GN HIFE.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Eagan / Serotype B;
RX MEDLINE=95012708; PubMed=7927773;
RA McCrea K.W., Watson W.J., Gilsdorf J.R., Marrs C.F.;
RT "Identification of hife and hife in the plus gene cluster of
RT Haemophilus influenzae type b strain Eagan.";
RL Infect. Immun. 62:4922-4928(1994).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=AM30 (770235) / Serotype B;
RX MEDLINE=95089703; PubMed=7997179;
RA Van Ham M.S., van Alphen L., Mool F.R., van Putten J.P.M.;
RT "The fibrial gene cluster of Haemophilus influenzae type b.";
RL Mol. Microbiol. 13:673-684(1994).
CC BIOGENESIS. MAY BE THE ADHESIVE COMPONENT FOR PILUS
CC -1- SIMILARITY: TO B.PERTUSSIS FIMD AND E.COLI FIMH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U13254; AAA61542.1; -
DR EMBL: Z33502; CAA83904.1; -
DR InterPro: IPR000259; Fimbrial.
DR Pfam: PF00419; Fimbrial; 1.

KW Fimbria; Signal.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 435 MINOR FIMBRIAL SUBUNIT HIFE.
SQ SEQUENCE 435 AA; 48851 MW; 420C733661DE7D67 CRC64;

Alignment Scores:
Pred. No.: 10.3 Length: 435
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x HFE1_HAEIN (1-435)
QY 776 CTTAAAAACACCTTCCCC 796
DB 389 LeuLysGlnProSerPro 395

RESULT 3
VCAP_HSVSA
ID VCAP_HSVSA STANDARD; PRT; 1371 AA.
AC 000999;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major capsid protein (MCP).
GN 25.
OS Herpesvirus saimiri (strain 11).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=10383;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92333688; PubMed=1321287;
RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,
RA Newman C., Wiltmann S., Craxton M.A., Coleman H., Fleckenstein B.,
RA Honess R.W.;
RT "Primary structure of the herpesvirus saimiri genome.";
RL J. Virol. 66:5047-5058(1992).
CC -1- FUNCTION: MAJOR PROTEIN OF THE ICOSAEDRAL CAPSID.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES MAJOR CAPSID PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X64346; CAA45648.1; -
DR PIR: E36808; VCHBM7.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP; 1.
DR PRINTS: PR00235; HSVCAPSIDMCP.
KW Coat protein; Late protein.
SQ SEQUENCE 1371 AA; 154354 MW; 5F99FA953A201C1 CRC64;

Alignment Scores:
Pred. No.: 8.15 Length: 1371
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x VCAP_HSVSA (1-1371)
QY 786 AACCTTCCGCCCTTTGCC 806
DB 560 AsnLeuProProLeuAla 566

```

RESULT 4
ID Y031_TREPA STANDARD; PRT; 92 AA.
AC 083074;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0031.
GN TP0031.
OS Treponema pallidum.
OC Bacteria: Spirochaetales: Spirochaetaceae: Treponema.
OX NCBI_TaxID=160;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Melnick G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AE001186; AAC65032.1; -
DR TIGR: TP0031; -
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 65 86 POTENTIAL.
SQ SEQUENCE 92 AA; 10276 MW; 4EB63881276C4E8 CRC64;

Alignment Scores:
Pred. No.: 195 Length: 92
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x Y031_TREPA (1-92)
OY 115 TTTTTCNAGCTTTT 98
DB 30 PhepneProtrpPhephe 35

RESULT 5
ID ING_CHICK STANDARD; PRT; 164 AA.
AC P49708;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon gamma precursor (IFN-gamma).
GN IFNG.
OS Gallus gallus (Chicken).
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
OC Archosauria: Aves: Neognathae: Galliformes: Phasianidae: Phasianinae:
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=96154849; PubMed=8590305;

```

```

RA Digby M.R., Lowenthal J.W.;
RT "Cloning and expression of the chicken interferon-gamma gene.";
RL J. Interferon Cytokine Res. 15:939-945(1995).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=97054661; PubMed=8898958;
RA Weinig K.C., Schultz U., Muenster U., Kaspers B., Staehel P.;
RT "Biological properties of recombinant chicken interferon-gamma.";
RL Eur. J. Immunol. 26:2440-2447(1996).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=N LINE;
RX MEDLINE=98172731; PubMed=9511740;
RA Kaiser P., Wain H.M., Rothwell L.;
RT "Structure of the chicken interferon-gamma gene, and comparison to
RT mammalian homologues.";
RL Gene 207:25-32(1998).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn; TISSUE=Spleen;
RA Song K.D., Kim D.K., Choi K.D., Lillehoj H.S., Han I.K., Han J.Y.;
RT "Molecular genetic analysis of chicken interferon gamma gene.";
RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC OF MACROPHAGES. IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC THE TYPE I INTERFERONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: U27465; AAC58723.1; -
DR EMBL: X95774; CAA68118.1; -
DR EMBL: Y07922; CAA69227.1; -
DR EMBL: U96875; AAG27918.1; -
DR EMBL: U96872; AAG27918.1; JOINED.
DR EMBL: U96873; AAG27918.1; JOINED.
DR EMBL: U96874; AAG27918.1; JOINED.
DR HSSP: P07353; ID9C.
DR InterPro: IPR002069; IFN-gamma.
DR Pfam: PF00714; IFN-gamma; 1.
DR ProDom: PD002435; IFN-gamma; 1.
KW Cytokine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 164 INTERFERON GAMMA.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 164 AA; 18953 MW; 18D9B0BCA415B600 CRC64;

Alignment Scores:
Pred. No.: 173 Length: 164
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x ING_CHICK (1-164)
OY 777 TTTAAACAAACCTGCC 794
DB 103 LeuYsAsnAsnLeuPro 108

```

```

RESULT 6
ING_MELGA
ID ING_MELGA STANDARD: PRT: 164 AA.
AC 057603;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon gamma precursor (IFN-gamma).
GN IFNG.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_Taxid=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Embryo;
RA Kaiser P., Sonnenmans D., Smith L.M.;
RT "Avian Interferon-gamma: cloning, sequencing and comparison of
RT Interferon-gamma genes from several different avian species.";
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC OF MACROPHAGES. IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC THE TYPE I INTERFERONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ000725; CA04260.1; -.
DR HSSP: P07353; ID9C.
DR InterPro: IPR002069; IFN-gamma.
DR Pfam: PF00714; IFN-gamma; 1.
DR ProDom: PD002435; IFN-gamma; 1.
DR CycloKine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 164 INTERFERON GAMMA.
FT CARBOHD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 164 AA; 18943 MW; 3488F5A9A05B5C7 CRC64;

Alignment Scores:
Pred. No.: 173 Length: 164
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.338 Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x ING_MELGA (1-164)
OY 777 TTAATAAACACCTCC 794
DB 103 LeuLysAsnAsnLeuPro 108

RESULT 7
ING_PHACO
ID ING_PHACO STANDARD: PRT: 164 AA.
AC 057608;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon gamma precursor (IFN-gamma).
GN IFNG.

```

```

OS Phasianus colchicus colchicus (Ring-necked pheasant).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Phasianus.
OX NCBI_Taxid=9057;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaiser P., Sonnenmans D., Smith L.M.;
RT "Avian Interferon-gamma: cloning, sequencing and comparison of
RT Interferon-gamma genes from several different avian species.";
RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PRODUCED BY LYMPHOCYTES ACTIVATED BY SPECIFIC ANTIGENS
CC OR MITOGENS. IFN-GAMMA, IN ADDITION TO HAVING ANTIVIRAL ACTIVITY,
CC HAS IMPORTANT IMMUNOREGULATORY FUNCTIONS. IT IS A POTENT ACTIVATOR
CC OF MACROPHAGES. IT HAS ANTIPROLIFERATIVE EFFECTS ON TRANSFORMED
CC CELLS AND IT CAN POTENTIATE THE ANTIVIRAL AND ANTITUMOR EFFECTS OF
CC THE TYPE I INTERFERONS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYPE II (OR GAMMA) INTERFERON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AJ001289; CA04649.1; -.
DR HSSP: P07353; ID9C.
DR InterPro: IPR002069; IFN-gamma.
DR Pfam: PF00714; IFN-gamma; 1.
DR ProDom: PD002435; IFN-gamma; 1.
DR CycloKine; Antiviral; Growth regulation; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 164 INTERFERON GAMMA.
FT CARBOHD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 164 AA; 18910 MW; 4B943A819405B716 CRC64;

Alignment Scores:
Pred. No.: 173 Length: 164
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.338 Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x ING_PHACO (1-164)
OY 777 TTAATAAACACCTCC 794
DB 103 LeuLysAsnAsnLeuPro 108

RESULT 8
RL9_CHLPN
ID RL9_CHLPN STANDARD: PRT: 169 AA.
AC 0926V3; Q9J051;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein l9.
GN RPL1 OR RL9 OR CP0953 OR CP0906.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_Taxid=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Gilmwood J., Davis R.W., Stephens R.S.;

```

```

RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39:
RX MEDLINE-20150255; PubMed-10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uettermann T., Berry K., Bass S.,
RA Linher K., Weldon J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gelin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mofn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-J138:
RX MEDLINE-20330349; PubMed-10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CML029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- FUNCTION: BINDS TO THE 23S RNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LSP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AEO01675; AAD19091.1; -
DR EMBL: AE002349; AAR38691.1; -
DR EMBL: AP002548; BAA99161.1; -
DR PHC1-2DPAGE: G926V3; -
DR TIGR: CP0906; -
DR InterPro: IPR000244; RIBOSOMAL_L9.
DR Pfam: PF01281; RIBOSOMAL_L9; 1.
DR TIGRfams: TIGR00158; L9; 1.
DR PROSITE: PS00651; RIBOSOMAL_L9; FAUSE-NEG.
KW RIBOSOMAL protein; RNA-binding; Complete proteome.
SQ SEQUENCE 169 AA; 18722 MW; A4A2E124F9EB4377 CRC64; .

Alignment Scores:
Pred. NO.: 172 Length: 169
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x RU9-CHLPM (1-169)
QY 99 AAAAAGCTGCAAAAAA 116
Db 124 Lyshenleuglylslys 129

RESULT 9
APT_MASHI STANDARD; PRT; 180 AA.
AC 064427;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
OS APRT.
GN Metonys hildebrandtii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mastomys.

```

```

OX NCBI_TaxID=34847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97176443; PubMed-9023989;
RA Fieldhouse D., Yazdani F., Golding G.B.;
RT "Substitution rate variation in closely related rodent species."
RL Heredity 78:21-31(1997).
CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC OF AMP, THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: AMP + diphosphate -> adenine + 5-phospho-
CC alpha-D-ribose 1-diphosphate.
CC -1- PATHWAY: Purine salvage.
CC -1- SUBUNIT: Purine salvage.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U28722; AAA68956.1; -
DR InterPro: IPR000836; PRTtransferase.
DR InterPro: IPR002375; Pr/py_rp.transf.
DR Pfam: PF00156; PRibosyltranf. 1.
DR TIGRfams: TIGR01090; apt; 1.
DR PROSITE: PS00103; PUR_PYR_PR_TRANSFER; 1.
KW Transferase; Glycosyltransferase; Purine salvage.
SQ SEQUENCE 180 AA; 19679 MW; 07E76B24159FF099 CRC64; .

Alignment Scores:
Pred. NO.: 170 Length: 180
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x APT_MASHI (1-180)
QY 771 GGACCATCCGCTTTTC 754
Db 169 Glyproleptrophephe 174

RESULT 10
APT_MOUSE STANDARD; PRT; 180 AA.
AC P08030;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
OS APRT.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-85190571; PubMed-3921964;
RA Dush M.K., Sikela J.M., Khan S.A., Tischfield J.A., Stambrook P.J.;
RT "Nucleotide sequence and organization of the mouse adenine
RT phosphoribosyltransferase gene: presence of a coding region common to
RT animal and bacterial phosphoribosyltransferases that has a variable
RT intron/exon arrangement."
RL Proc. Natl. Acad. Sci. U.S.A. 82:2731-2735(1985).
CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC OF AMP, THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: AMP + diphosphate -> adenine + 5-phospho-

```

```
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M1310; AAA37255.1; -.
DR      PIR; A22670; RTMSA.
DR      MGI; 88061; Aprt.
DR      InterPro: IPR000836; PRTtransferase.
DR      InterPro: IPR002375; Pr/PY-TP-Transf.
DR      Pfam; PF00156; Pribosylltran; 1.
DR      TIGRPFAMS; TIGR01090; apt; 1.
DR      PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
DR      TRANSFERASE; Glycosyltransferase; Purine salvage.
KW      TRANSFERASE
SQ      SEQUENCE 180 AA; 19736 MW; 8EAF0CB4F173C64 CRC64;

Alignment Scores:
Pred. No.: 170          Length: 180
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 1                  Gaps: 0

US-09-835-992a-22 (1-843) x APT_MOUSE (1-180)
OY      771 GGACGATCCCGTTTTC 754
DB      169 GYPRDLEPHEPHE 174

RESULT 11
APT_MOUSE
ID      APT_MOUSE          STANDARD:      PRT: 180 AA.
AC      P47956;
DR      01-FEB-1996 (Rel. 33, Created)
DR      01-FEB-1996 (Rel. 33, Last sequence update)
DR      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN      APRT.
OS      Mus pahari (Shrew mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10093;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97176443; PubMed=9023989;
RA      Fieldhouse D., Yazdani F., Golding G.B.;
RT      "Substitution rate variation in closely related rodent species.";
RL      Heredity 78:21-31(1997).
CC      -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC      OF AMP. THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC      -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
```

```
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U28721; AAA68957.1; -.
DR      InterPro: IPR000836; PRTtransferase.
DR      InterPro: IPR002375; Pr/PY-TP-Transf.
DR      Pfam; PF00156; Pribosylltran; 1.
DR      TIGRPFAMS; TIGR01090; apt; 1.
DR      PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
DR      TRANSFERASE; Glycosyltransferase; Purine salvage.
KW      TRANSFERASE
SQ      SEQUENCE 180 AA; 19712 MW; A8F43EB56F4D2CBF CRC64;

Alignment Scores:
Pred. No.: 170          Length: 180
Score: 6.00            Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.28%      Indels: 0
DB: 1                  Gaps: 0

US-09-835-992a-22 (1-843) x APT_MOUSE (1-180)
OY      771 GGACGATCCCGTTTTC 754
DB      169 GYPRDLEPHEPHE 174

RESULT 12
APT_MOUSE
ID      APT_MOUSE          STANDARD:      PRT: 180 AA.
AC      P47957;
DR      01-FEB-1996 (Rel. 33, Created)
DR      01-FEB-1996 (Rel. 33, Last sequence update)
DR      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Adenine phosphoribosyltransferase (EC 2.4.2.7) (APRT).
GN      APRT.
OS      Mus spicilegus (Steppe mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10103;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97176443; PubMed=9023989;
RA      Fieldhouse D., Yazdani F., Golding G.B.;
RT      "Substitution rate variation in closely related rodent species.";
RL      Heredity 78:21-31(1997).
CC      -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
CC      OF AMP. THAT IS ENERGETICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
CC      -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
CC      alpha-D-ribose 1-diphosphate.
CC      -1- PATHWAY: Purine salvage.
CC      -1- SUBUNIT: HOMODIMER.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC      -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
CC      PHOSPHORIBOSYLTRANSFERASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U28720; AAA68958.1; -.
DR      InterPro: IPR000836; PRTtransferase.
DR      InterPro: IPR002375; Pr/PY-TP-Transf.
DR      Pfam; PF00156; Pribosylltran; 1.
DR      TIGRPFAMS; TIGR01090; apt; 1.
DR      PROSITE; PS00103; PUR_PYR_PR_TRANSFER; 1.
DR      TRANSFERASE; Glycosyltransferase; Purine salvage.
KW      TRANSFERASE
SQ      SEQUENCE 180 AA; 19724 MW; A6AE4DBEDB45E42 CRC64;
```


Alignment Scores:

Prod. No.:	Score:	Length:	Matches:
170	6.00	180	6
Percent Similarity:	100.00%		Conservative: 0
Best Local Similarity:	100.00%		Mismatches: 0
Query Match:	3.28%		Indels: 0
DB:	1		Gaps: 0

US-09-835-992a-22 (1-843) x APT_MUSSI (1-180)

OY 771 GGACNATCCCGTTTC 754
 Db 169 GTPProlleProphene 174

RESULT 13
 RRF_DEIRA STANDARD: PRT: 183 AA.

AC 09R082;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ribosome recycling factor (Ribosome releasing factor) (RRF).
 GN RRF OR DR1510.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RI;
 RA MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffitt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Makarewicz J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioreistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 CC -1- FUNCTION: Responsible for the release of ribosomes from messenger
 CC RNA at the termination of protein biosynthesis. May increase the
 CC efficiency of translation by recycling ribosomes from one round of
 CC translocation to another (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE RRF FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AEO01995; AAF1077.1; -
 CC DR HSSP: O9X189; 1DD5.
 CC DR TIGR: DR1510; -
 CC DR InterPro: IPR002661; RRF.
 CC DR Pfam: PF01765; RRF; 1.
 CC DR TIGRPFAM: TIGR00496; frr; 1.
 CC KW Protein biosynthesis; Complete proteome.
 CC SQ SEQUENCE 183 AA; 20620 MW; 40B1546728CE5B57 CRC64;

Alignment Scores:

Prod. No.:	Score:	Length:	Matches:
170	6.00	183	6
Percent Similarity:	100.00%		Conservative: 0
Best Local Similarity:	100.00%		Mismatches: 0
Query Match:	3.33%		Indels: 0
DB:	1		Gaps: 0

US-09-835-992a-22 (1-843) x RRF_DEIRA (1-183)

OY 823 CTAAATCCGACATATA 840
 Db 91 LeuansProAnaAnlys 96

RESULT 14
 YPDC_BACSU STANDARD: PRT: 218 AA.

ID YPDC_BACSU
 AC P50738;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein ypdC.
 GN ypdC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168 / Marburg;
 RX MEDLINE=96349105; PubMed=8760912;
 RA Sorokin A.V., Azevedo V., Zumbstein E., Galleron N., Ehrlich S.D.,
 RA Serror P.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RT the sera and kds loci cloned in a yeast artificial chromosome.";
 RL Microbiology 142:2005-2016(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berrtero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriss L., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghm S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
 RA Pressecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schreier R., Scoffone F.,
 RA Seliguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Solido B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yamaoka K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: LA7648; AAC83956.1; -

DR EMBL: 299115; CAB14210.1; -
DR EMBL: 299116; CAB14226.1; -
DR Subtilist; Bg11438; yPdc.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 30 50 POTENTIAL.
FT TRANSMEM 65 85 POTENTIAL.
FT TRANSMEM 127 147 POTENTIAL.
FT TRANSMEM 180 200 POTENTIAL.
SQ SEQUENCE 218 AA; 24719 MW; 88556D50863E14BC CRC64;

Alignment Scores:
Pred. No.: 164 Length: 218
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x YPDC_BACSU (1-218)

QY 676 TTTTACCTTGCTAAGGCN 693
Db 147 PhetYrLeuGlYsAla 152

RESULT 15
YRAL_YEAST STANDARD; PRT; 226 AA.
AC 012159;
DR 15-JUL-1998 (Rel. 36, Created)
DR 15-JUL-1998 (Rel. 36, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA annealing protein YRAL.
GN YRAL OR YDR381W OR D9481.2 OR D9509.1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX NCBI-TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 39-53 AND 57-70.
RC STRAIN-BJ5626;
RX MEDLINE=97293180; PubMed=9149233;
RA Portman D.S., O'Connor J.P., Dreyfuss G.;
RT "YRAL, an essential Saccharomycos cerevisiae gene, encodes a novel
RT nuclear protein with RNA annealing activity.";
RL RNA 3:527-537(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288c / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favella A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
RA Talch A., Trevaslis E., Vignati D., Wilcox L., Woldman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
RA Bero A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunnicke-Smith S., Hyman R., Komp C., Laishari D., Lew H., Lin D.,
RA Moseedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RNA-BINDING RNA ANNEALING PROTEIN. MAY HAVE A ROLE IN
CC PRE-MRNA METABOLISM.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -!- SIMILARITY: TO S.POMBE MJO3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U72633; AAC09951.1; -
DR EMBL: U28373; AAB64817.1; -
DR EMBL: U32274; AAB64823.1; -
DR SGD: S0002789; YRAL.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PR00076; rrm; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding.
FT DOMAIN 78 158 RNA-BINDING (RRM).
FT DOMAIN 201 210 ARG/YXS-RICH (BASIS).
SQ SEQUENCE 226 AA; 24955 MW; 480B5B6DC0D14BE9 CRC64;

Alignment Scores:
Pred. No.: 162 Length: 226
Score: 6.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-09-835-992a-22 (1-843) x YRAL_YEAST (1-226)

QY 619 TTTACGGGTCGCCAAT 636
Db 139 PhasnglySerProtle 144

Search completed: January 14, 2003, 17:29:21
Job time: 17.0173 secs

SQ SEQUENCE 288 AA; 33108 MW; 06441683774F7B14 CRC64;
 Alignment Scores:
 Pred. No.: 1.86 Length: 288
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.37% Indels: 0
 DB: 5 Gaps: 0
 US-09-835-992a-22 (1-843) x Q95011 (1-288)
 QY 127 TTTGCAATATTTTTCNCNAG 104
 DB 131 PheAlaIleIlePheProArg 138
 RESULT 2
 Q41572 PRELIMINARY; PRT; 86 AA.
 AC Q41572:
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE (Begra) ORP1 and ORP2 genes.
 OS Triticum aestivum (wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. BEGRA;
 RA Kroczyńska B., Buchowicz J.,
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X76181; CAAS3774.1; -.
 SQ SEQUENCE 86 AA; 9783 MW; 82D92DDBD138242 CRC64;
 Alignment Scores:
 Pred. No.: 37.6 Length: 86
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.83% Indels: 0
 DB: 10 Gaps: 0
 US-09-835-992a-22 (1-843) x Q41572 (1-86)
 QY 799 GGNGGGGAAGTGTGTTTTT 779
 DB 37 GlyGlyGlyArgLeuPhePhe 43
 RESULT 3
 Q53177 PRELIMINARY; PRT; 172 AA.
 AC Q53177:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein RV2450c.
 GN RV2450c OR MYV008.06c.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RA MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felkell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RA complete genome sequence."
 RL Nature 393:537-544 (1998).
 DR EMBL: AL021246; CAI16027.1; -.
 DR Tuberculosis; RV2450c; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 172 AA; 17456 MW; 0D67E72240EFOF0E CRC64;
 Alignment Scores:
 Pred. No.: 33.2 Length: 172
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 16 Gaps: 0
 US-09-835-992a-22 (1-843) x Q53177 (1-172)
 QY 786 AACCTTCCCCNCCTTGGCC 806
 DB 73 AsnLeuProProLeuAla 95
 RESULT 4
 Q8VJ14 PRELIMINARY; PRT; 188 AA.
 AC Q8VJ14:
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein MT2526.
 GN MT2526.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE007090; AAK46825.1; -.
 DR TIGR: MT2526; -.
 KW Hypothetical protein.
 SQ SEQUENCE 188 AA; 19223 MW; 3AF2AD19B97736E2 CRC64;
 Alignment Scores:
 Pred. No.: 32.6 Length: 188
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 16 Gaps: 0
 US-09-835-992a-22 (1-843) x Q8VJ14 (1-188)
 QY 786 AACCTTCCCCNCCTTGGCC 806
 DB 89 AsnLeuProProLeuAla 95
 RESULT 5
 Q9BL14 PRELIMINARY; PRT; 215 AA.
 AC Q9BL14:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE Fibroin P25.
 GN FIB-P25.
 OS Papilio xuthus.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Papilionidae; Papilioninae; Papilio.
 OC NCBITaxID=66420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21167527; PubMed=11267905;
 RA Tanaka K., Mizuno S.;
 RT "Homologues of fibroin L-chain and P25 of Bombyx mori are present in
 RT Dendrolimus spectabilis and Papilio xuthus but not detectable in
 RT Anthrenus yamamai.";
 RL Insect Biochem. Mol. Biol. 31:665-677(2001).
 DR EMBL; AF001825; BAB39504.1; -. 21BD9C93FDF83389 CRC64;
 SQ SEQUENCE 215 AA; 24346 MW; 21BD9C93FDF83389 CRC64;

Alignment Scores:
 Pred. No.: 31.8 Length: 215
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 5 Gaps: 0

US-09-835-992a-22 (1-843) x Q9BLU4 (1-215)

OY 820 CGTCTAATCCGACATATA 840

DB 165 ArgLeuAsnProAsnAsnLys 171

RESULT 6
 Q98S75 PRELIMINARY; PRT; 223 AA.

AC Q98S75;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, last annotation update)
 DE Hypothetical 26.9 kDa protein of f223 from chromosome 3.
 GN ORF223.
 OS Guillardia theta (Cryptomonas phl).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBITaxID=55529;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21223349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Mäler U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39707.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 223 AA; 26902 MW; A0396B214866F092 CRC64;

Alignment Scores:
 Pred. No.: 31.6 Length: 223
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x Q98S75 (1-223)

OY 99 AAAACCTGCAAAAAATA 119

DB 87 LysAsnLeuGlyLysLysIle 93

RESULT 7
 Q9FR36

ID Q9FR36 PRELIMINARY; PRT; 263 AA.
 AC Q9FR36;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Putative steroid reductase.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseolaceae; Glycine.
 OC NCBITaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, WILLIAMS 82; TISSUE=LEAF;
 RC MEDLINE=99169286; PubMed=10069834;
 RA Reverdatto S., Beilinson V., Nielsen N.C.;
 RT "A multisubunit acetyl coenzyme A carboxylase from soybean.";
 RL Plant Physiol. 119:961-978(1999).
 DR EMBL; AF203341; AAG35638.1; -.
 DR InterPro; IPR001104; Strd5A_dhc.
 DR Pfam; PF02544; Steroid_dh; 1.
 SQ SEQUENCE 263 AA; 30263 MW; 22DC8E76D501135 CRC64;

Alignment Scores:
 Pred. No.: 30.7 Length: 263
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x Q9FR36 (1-263)

OY 786 AACCTCCGCCCTTGCC 806

DB 48 AsnLeuProProLeuAla 54

RESULT 8
 Q9FSH2 PRELIMINARY; PRT; 328 AA.

AC Q9FSH2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
 DE Riorf20 protein.
 GN RIORF20.
 OS Agrobacterium rhizogenes.
 OG Plasmid pRI1724.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OC NCBITaxID=359;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MAF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Satuli N., Kataoka M., Tanaka N.,
 RA Yoshida K.;
 RT "The complete nucleotide sequence of a R1 (root inducing) plasmid
 RT indicates its chimerical structure between T1 and Sym plasmids.";
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MAF03-01724;
 RA Moriguchi K., Maeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;
 RT "Analysis of unique variable region of a plant root inducing plasmid,
 RT pRI1724, by the construction of its physical map and library.";
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=MAF03-01724;
 RA Moriguchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;

RT "Genome structure of R1 plasmid (1): Construction of linking library
 RT and physical map of pRI1724 in Japanese Agrobacterium.";
 RL Nucleic Acids Symp. Ser. 39:189-190(1998).

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAEP03-01724;
RX MEDLINE=20241294; PubMed=10780382;
RA Maeda Y., Moriguchi K., Kataoka M., Satou M., Satuli N., Tanaka N.,
RT Yoshida K.:
RT "Genome structure of R1 plasmid (1): Sequencing analysis of T-DNA and
RT its flanking regions of pR11724 in Japanese Agrobacterium
RT Rhizogenes.";
RL Nucleic Acids Symp. Ser. 42:67-68(1999).
DR EMBL; AP002086; BAB16139.1; -.
RN Plasmid.
SQ SEQUENCE 328 AA; 36666 MW; 0632204AA307C2A5 CRC64;

Alignment Scores:
Pred. No.: 29.5 Length: 328
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: Gaps: 0

US-09-835-992a-22 (1-843) x Q9F5H2 (1-328)

QY 673 CCATTTCCTGCTGAAGGCN 693
DB 304 PProPhetYrLeugLYsAla 310

RESULT 9
003951 PRELIMINARY; PRT; 387 AA.
ID 003951.
AC 003951.
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DR 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GP28 (Partition protein A).
GN GENE 28 OR PARA.
OS Bacteriophage N15.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=40631;
RN [1]
RP SEQUENCE FROM N.A.
RA Hendrix R.W., Ravin V.K., Casjens S.R., Ford M.E., Ravin N.V.,
RL Smirnov I.K.:
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-207 FROM N.A.
RA Svarchevsky A.N., Rybchin V.N.:
RL Mol. Gen. Microbiol. Virusol. 2:16-22(1984).
RN [3]
RP SEQUENCE OF 1-207 FROM N.A.
RA Vostrov A.A., In Malinin A., Rybchin V.N., Svarchevsky A.N.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF064539; AAC19064.1; -.
DR EMBL; U91583; AAB58384.1; -.
DR InterPro; IPR000707; ATPase_PARA.
DR Pfam; PF00991; Para; 1.
SQ SEQUENCE 387 AA; 43607 MW; 40FD9058078F992F CRC64;

Alignment Scores:
Pred. No.: 28.6 Length: 387
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: Gaps: 0

US-09-835-992a-22 (1-843) x 003951 (1-387)

QY 670 TTACATTTTACCTGCTAAG 690
DB 169 LeuProPhetYrLeugLYs 175
```

```

RESULT 10
09GKN7 PRELIMINARY; PRT; 426 AA.
ID 09GKN7.
AC 09GKN7.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Rh50-like protein.
GN Rh50.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheriff-Zahar B., Raynal V., Cartlon J.P., Matassi G.:
RT "Evolution of Rh50 genes in Metazoa.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164575; AAG38955.1; -.
DR InterPro; IPR001905; Ammonium_transp.
DR InterPro; IPR002229; RhesusRHD.
DR InterPro; IPR001693; SHProt_acsite.
DR Pfam; PF00909; Ammonium_transp; 1.
DR PRINTS; PR00342; RHEUSRHD.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; UNKNOWN_1.
FT VARIANT 201 201 S -> P.
SQ SEQUENCE 426 AA; 46719 MW; C4899A8406E09E51 CRC64;

Alignment Scores:
Pred. No.: 28.1 Length: 426
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.89% Indels: 0
DB: Gaps: 0

US-09-835-992a-22 (1-843) x Q9GKN7 (1-426)

QY 409 TTTTAATTTTAAACCCCTA 429
DB 313 PheLYsPheLeuThrProLeu 319

RESULT 11
099Q04 PRELIMINARY; PRT; 435 AA.
ID 099Q04.
AC 099Q04.
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DR 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE H1FE.
GN H1FE.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=1595, 6351, AND 785;
RC Clemons D.L., Patel M.T., Harris C.F., Bauer R.J., Gilstorf J.R.;
RT "Analysis of the pilus adhesins from Haemophilus influenzae biotype IV
RT strains.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245361; AAK28377.1; -.
DR EMBL; AF245359; AAK28375.1; -.
DR EMBL; AF245360; AAK28376.1; -.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
SQ SEQUENCE 435 AA; 48807 MW; E2E157E3E018677A CRC64;

Alignment Scores:
Pred. No.: 28 Length: 435
```

Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x Q99Q04 (1-435)

OY 776 CTTAAACCAACCTTCCCCC 796
 DB 389 LeuLysLysGlnProSerPro 395

RESULT 12

O9AIR2 ID O9AIR2 PRELIMINARY: PRT: 435 AA.

AC O9AIR2; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE H1FE.
 CN H1FE.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus; 727;
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SFRAN-1216;
 RA Clemons D.L., Patel M.J., Marrs C.F., Bauer R.J., Glisford J.R.;
 RT "Analysis of the pilus adhesins from Haemophilus influenzae biotype IV
 RT strains."
 RL EMBL: AF245363; AAK8379.1;
 DR InterPro: IPR000259; Fimbril.
 DR Pfam: PF00419; Fimbril; 1.
 SQ SEQUENCE 435 AA; 48767 MW; 477DBE92A749CEAD CRC64;

Alignment Scores:

Pred. No.: 28 Length: 435
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 2 Gaps: 0

US-09-835-992a-22 (1-843) x O9AIR2 (1-435)

OY 776 CTTAAACCAACCTTCCCCC 796

DB 389 LeuLysLysGlnProSerPro 395

RESULT 13

O9M2D7 ID O9M2D7 PRELIMINARY: PRT: 498 AA.

AC O9M2D7; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE Hypothetical 58.5 kDa protein.
 CN T20K12.170.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

AC O9M2D7; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE Hypothetical 58.5 kDa protein.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

AC O9M2D7; 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE Hypothetical 58.5 kDa protein.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL137898; CAB71057.1;
 KW Hypothetical protein.
 SQ SEQUENCE 498 AA; 58523 MW; B2E8606256BB8E4A CRC64;

Alignment Scores:

Pred. No.: 27.3 Length: 498
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 10 Gaps: 0

US-09-835-992a-22 (1-843) x Q9M2D7 (1-498)

OY 817 TCCGCTTAATCCGACCAAT 837
 DB 84 SerArgLeuAsnProAsnAsn 90

RESULT 14

O9P6U5 ID O9P6U5 PRELIMINARY: PRT: 1240 AA.

AC O9P6U5; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Related to protease ULP2 protein.
 GN 15E6.80.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algn V., Hohelsel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Meves H.W., Mannhaupt G.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL353822; CAB8639.1;
 DR InterPro: IPR003653; SUMO_protease.
 DR Pfam: PF02902; Peptidase_C48; 1.
 KW Protease.
 SQ SEQUENCE 1240 AA; 138114 MW; 716E38F4DF0D177A CRC64;

Alignment Scores:

Pred. No.: 23.1 Length: 1240
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.89% Indels: 0
 DB: 3 Gaps: 0

US-09-835-992a-22 (1-843) x Q9P6U5 (1-1240)

OY 767 GGTCCGCCCTTAAACCA 787

DB 29 GlyProLeuLysLysGln 35

RESULT 15

O9YTP1 ID O9YTP1 PRELIMINARY: PRT: 1371 AA.

AC O9YTP1; 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Major capsid protein.
 OS Ateline herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=85618;

AC O9YTP1; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DE Major capsid protein.

OS Ateline herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=85618;

```

RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=73;
RX      MEDLINE=20091363; PubMed-10623770;
RA      Albrecht J.C.;
RT      "Primary structure of the Herpesvirus Ateles genome.";
RL      J. Virol. 74:1033-1037(2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=73;
RA      Albrecht J.-C.; Fleckenstein B.;
RL      Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF083424; AAC95549.1; -
DR      InterPro; IPR000912; Herpes_MCP.
DR      Pfam; PF03122; Herpes_MCP; 1.
DR      PRINTS; PR00235; HSCAPSIDMCP.
SQ      SEQUENCE 1371 AA; 153800 MW;  EE9FC518FDF066A0 CRC64;

Alignment Scores:
Pred. No.:      22.7      Length:      1371
Score:          7.00      Matches:      7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    3.89%      Indels:      0
DB:             12      Gaps:        0

US-09-835-992a-22 (1-843) x O9YTP1 (1-1371)
QY      786 AACCTTCCCCCNCCTTTGGCC 806
        |||||
Db      560 AsnLeuProProProLeuAla 566
```

Search completed: January 14, 2003, 17:37:09
Job time : 65.3589 secs